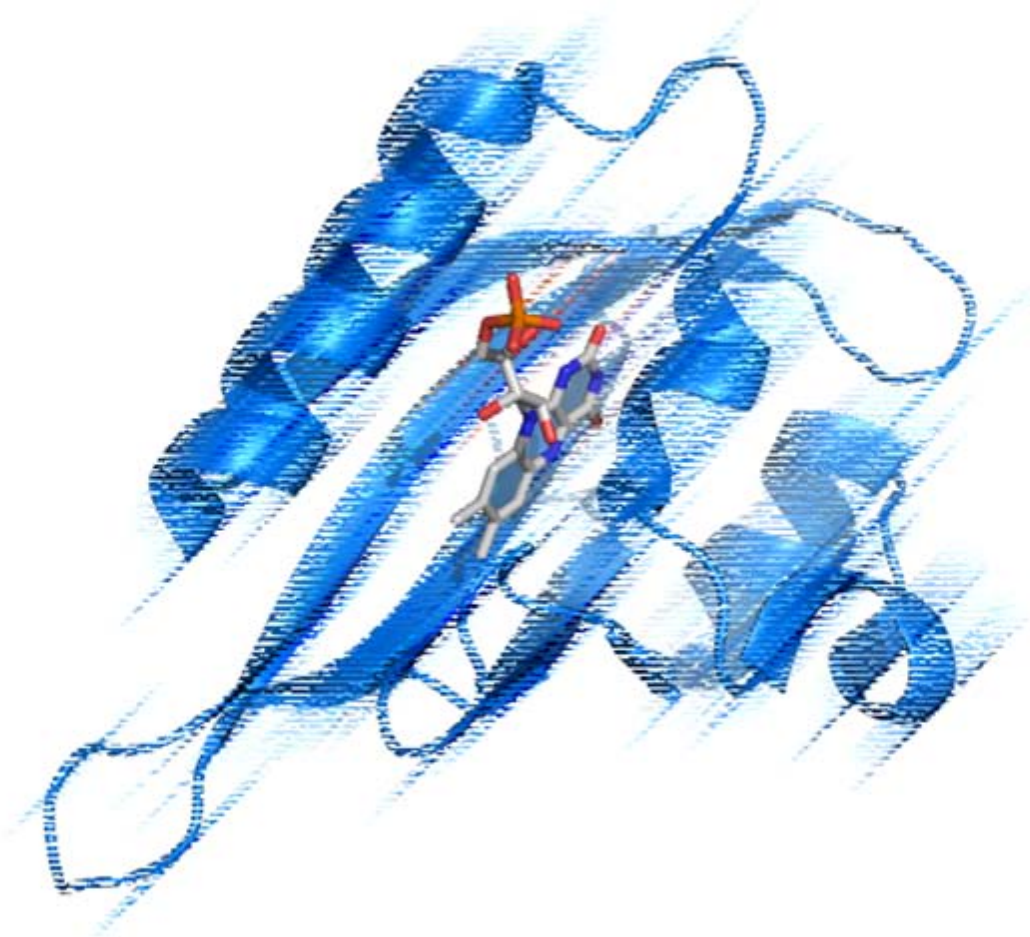
Bacterial Blue-Light Photoreceptors of the LOV Family



Ulrich Krauss

Bacterial Blue-Light Photoreceptors of the LOV Family

Inaugural – Dissertation

Zur
Erlangung des Doktorgrades der
Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine Universität Düsseldorf

vorgelegt von
Diplom-Naturwissenschaftler
Ulrich Krauss

aus Marienberg/Sachsen

November 2007

Aus dem Institut für Molekulare Enzymtechnologie

der Heinrich-Heine Universität Düsseldorf

Gedruckt mit Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. K.-E. Jäger

Koreferent: Prof. Dr. W. Gärtner

Tag der mündlichen Prüfung: 14.01.2008

dedicated to the two women in my life: Karen and Luna

“To see a world in a grain of sand,

And heaven in a wild flower,

Hold infinity in the palms of your hand,

And eternity in an hour”

WILLIAM BLAKE, Auguries of Innocence

I) Summary

In 1998, Briggs and co-workers identified the so-called light, oxygen, voltage (LOV) domains as the blue-light sensitive flavin-binding signaling switches in plant phototropins (phots) [1], controlling plant phototropism and other blue-light dependent phenomena [2]. It soon became apparent that many photosynthetic, but also non-photosynthetic prokaryotes possess homologous proteins encoded in their respective genomes [3]. Simultaneously with this discovery in 2002, their photochemical functionality was proven by Losi and co-workers [3] exemplary for the LOV domain-containing protein YtvA from *Bacillus subtilis*. This led to the suggestion that the LOV signaling paradigm might be conserved between Pro- and Eukaryotes. Mechanistically, the light sensing function of the LOV proteins is strictly based on the presence of a cysteine residue located in a distance of about 4 Å to the isoalloxazine ring of the flavin chromophore. Upon irradiation, a covalent bond is formed between this cysteine and the carbon-atom in position 4a of the flavin isoalloxazine ring, which thermally opens again within minutes to hours, dependent on the LOV protein [4, 5].
With this thesis the available knowledge about bacterial LOV proteins was broadened with respect to a) evolutionary history and b) signal-transduction and photophysiological aspects. Furthermore, some of the newly described bacterial LOV proteins were optimized with regard to their c) biotechnological application as fluorescent markers.

a) Evolutionary history of LOV proteins

A comprehensive phylogenetic analysis was performed that included all currently identified prokaryotic LOV protein/gene sequences as well as representatives from the three major eukaryotic LOV photoreceptor families, namely plant phot-LOVs, ZEITLUPE(ZTL)/ ADAGIO(ADO)/ Flavin-Binding Kelch-Repeat F-Box protein (FKF1)-LOVs and the fungal white-collar-1 (WC-1)-LOVs. The analysis suggested that the ancient LOV signaling module, that apparently retained its photosenitivity from Archaea to plants, spread from the prokaryotic to the eukaryotic kingdom of life by two independent endosymbiotic events. The plant phot-LOVs as well as the fungal WC-1 LOVs show clear affinity towards the α-proteobacterial clade of the phylogentic tree. Therefore, they are construed to originate from the endosymbiosis of an ancient proteobacterium that also led to the appearance of the mitochondrion in eukaryotes. On the other hand the plant ZTL/ADO/FKF1-LOV domains clearly cluster within the cyanobacteria, and thus endosymbiosis of an ancient cyanobacterium should mark the appearance of this LOV photoreceptor family (and moreover coincides with the appearance of the chloroplasts) in the eukaryotic kingdom.
Additionally, prokaryotic LOV histidine kinases are assumed as the primordial LOV photoreceptor systems in all three kingdoms of life, with a considerable amount of domain

shuffling (fusion and fission) implied in the process of full-length LOV photoreceptor evolution.

b) Blue-light dependent LOV signal-transduction and physiology in bacteria

Two novel LOV photoreceptor modules of the saprotrophic Pseudomonad *Pseudomonas putida* KT2440 were identified, cloned, expressed and purified from *Escherichia coli* as heterologous host. The two proteins named PpSB1-LOV and PpSB2-LOV both bind oxidized flavin mononucleotide as chromophore and show a phot-like primary photochemistry. However, although the two paralogous proteins share about 66% identical positions on the amino acid level, they exhibit dramatically different dark recovery kinetics. PpSB1-LOV is the slowest ever described bacterial LOV protein having an apparent recovery time of about 30 hours at 20°C, while PpSB2-LOV, on the other hand, shows a remarkable fast recovery of about 100 seconds at 20°C. Both proteins possess, apart from the conserved LOV core, only short N- and C-terminal extensions but completely lack a fused effector domain. A bioinformatic analysis of the genomic context in which the duplicated LOV genes are found in *P. putida* revealed a clustering of the LOV genes together with genes known and annotated to be involved in the iron-starvation response in this organism. Since in bacteria functionally related genes are often clustered together in large gene regions the involvement of the blue-light receptors in the regulation of iron-uptake has been investigated. The corresponding photophysiological experiments could demonstrate that blue-light preferentially enhances the secretion of the iron-siderophore pyoverdine, in *P. putida* (wildtype) culture supernatants, when grown under iron-limitation. Furthermore, the same effect is absent when the strain is grown under red- and green-light, thus suggesting the involvement of a blue-light photoreceptor in this response.

In addition, comparative analysis of secondary structure (applying circular dichroism spectroscopy) and quaternary structure (using size-exclusion chromatography) of YtvA and its isolated LOV domain, as well as PpSB1-LOV and PpSB2-LOV, highlighted the importance of dimerization of bacterial LOV domains and proteins and could furthermore, together with computational docking simulations, suggest a common surface, namely the central β-scaffold of the core LOV domain, for the LOV-LOV or LOV-Effector interaction in *B. subtilis* YtvA. This information together with previous experimental evidence that indicated similar processes in eukaryotic phot-LOVs [6-8] point towards a common interaction surface that could be involved in the signal-transduction process, both in the pro- and eukaryotic LOV photoreceptors.

c) Application of bacterial LOV proteins

Some already characterized bacterial LOV proteins (YtvA and its isolated LOV domain) and one of the newly identified *P. putida* LOV proteins (PpSB2-LOV) were further mutationally optimized for the use as autofluorescent marker proteins. The photoactive cysteine residue (highlighted in bold) of the LOV domain's canonical sequence motif GXN**C**RFLQG was mutationally replaced by a non-polar alanine residue. This mutation abolishes the photocycle in the LOV domain, resulting upon excitation with blue-light, in a continous switching of the LOV domain flavin co-factor between the ground and its excited singlet-state. Radiative decay of the singlet to the ground-state is accompanied by the emission of photons in the form of fluorescence. As this process is independent of the availability of molecular oxygen, the newly generated autofluorescent proteins, unlike the Green Fluorescent Protein (GFP) and its variants (which depend on molecular oxygen for the cyclization of their fluorescent chormophores), should allow non-invasive fluorescence imaging in anoxygenic microorganisms and oxygen-deprived zones of cellular tissues. Therefore, the newly generated FMN-dependent fluorescent proteins (FbFPs) which originate from bacterial LOV photoreceptor proteins, tackle the one major drawback of GFP fluorescent reporters and should thus in the future allow the study of currently difficult to analyze biological settings.

II) Zusammenfassung

Im Jahr 1998 identifizierte die Arbeitsgruppe um Winslow Briggs sogenannte *Light, Oxygen, Voltage* (LOV) Domänen als die Blaulicht-sensitiven Sensor Module in pflanzlichen Phototropinen (phots) [1]. Diese bedingen bestimmte physiologische Effekte, wie zum Beispiel den pflanzlichen Phototropismus, sowie andere Blaulicht-abhängige Phänomene wie die Chloroplastenbewegung und Blatt-Öffnung [2]. Weithin wurde sehr schnell klar, dass sehr viele photosynthetische sowie viele nicht-photosynthetisch aktive Prokaryoten ebenfalls homologe Proteine besitzen [3]. Mit dieser Entdeckung, die Losi, Gärtner *et al.* [3] im Jahr 2002 gelang, konnte gleichzeitig exemplarisch die photochemische Funktionalität prokaryotischer LOV-Proteine am Beispiel des aus *Bacillus subtilis* stammenden YtvA-Proteins nachgewiesen werden. Dies führte wiederum zu der Annahme, dass die LOV Sensor-Funktion sowohl im Prokaryoten- als auch im Eukaryoten-Reich strukturell und funktionell konserviert ist.

Mechanistisch beruht die LOV Sensor-Funktion auf der Ausbildung einer kovalenten Bindung zwischen einem hochkonservierten Cysteinrest in der LOV-Domäne und dem C4a-Kohlenstoffatom des Flavin-Kofaktors. Diese Bindung wird im Dunkeln thermisch innerhalb von Minuten bis Stunden wieder aufgelöst, wonach die LOV-Domäne wieder in ihren Grundzustand zurückkehrt.

Ziel der vorliegenden Arbeit war es, das vorhandene Wissen über bakterielle LOV-Proteine im Allgemeinen zu erweitern, wobei die Aspekte a) Evolution und b) Signalweiterleitung sowie damit verbundene photophysiologische Effekte untersucht wurden. Weiterhin konnten einige der hier neu beschriebenen bakteriellen LOV-Proteine erfolgreich für c) die biotechnologische Anwendung als Fluoreszenz-Marker optimiert werden.

a) Evolution der LOV Sensor-Funktion

Im Rahmen dieser Arbeit wurde eine umfangreiche phylogenetische Analyse, basierend auf allen momentan identifizierbaren prokaryotischen LOV-Gensequenzen sowie auf einer Auswahl an Sequenzen der drei größten eukaryotischen LOV-Photorezeptorfamilien (pflanzliche phot-LOV's und ZEITLUPE(ZTL)/ ADAGIO(ADO)/ Flavin-Binding Kelch-Repeat F-Box Protein (FKF1)-LOV's sowie die White-Collar-1 LOV-Proteine aus der Familie der Pilze), durchgeführt. Die Analyse legt nahe, dass das LOV Sensormodul sich zweimal unabhängig voneinander über den bakteriellen Endosymbionten, von den Prokaryoten auf Eukaryoten verbreitet hat. Die Gruppe der pflanzlichen Phototropin LOV-Domänen sowie die Gruppe der White-Collar-1-verwandten Proteine, die ausschließlich in Pilzen vorkommen, zeigen eine klare Affinität zu bestimmten α-proteobakteriellen Zweigen des phylogenetischen Stammbaums. Die LOV-Domänen, die als Sensormodule in den entsprechenden

eukaryotischen Photorezeptoren Einsatz finden sollten, waren also höchstwahrscheinlich schon im proteobakteriellen Vorfahren des eukaryotischen Endosymbionten, der heute das Mitochondrium bildet, vorhanden. Die dritte eukaryotische LOV-Photorezeptorfamilie (ZTL/ADO/FKF1-LOV) ist offensichtlich direkt mit bestimmten LOV-Sequenzen aus Cyanobakterien verwandt. Demzufolge sollte auch hier Endosymbiose (eines Cyanobakteriums als Proto-Chloroplast) für das Auftauchen der LOV-Photorezeptoren in den Eukaryoten verantwortlich sein. Zusätzlich geben die durchgeführten Analysen Grund zu der Annahme, dass bakterielle LOV-Histidin-Kinasen zu den primordialen LOV-Sensor-Systemen gehören könnten. Außerdem muss ein beachtliches Maß an Domänen-Fusionen und Abspaltungen stattgefunden haben, damit die enorme Vielfalt an mit LOV-fusionierten Effektor-Domänen erklärt werden kann.

b) Blaulicht-abhängige Signalweiterleitung und photophysiologische Effekte in Bakterien

Zwei putative LOV-Photorezeptoren aus *Pseudomonas putida* KT2440 konnten im Rahmen dieser Arbeit identifiziert, kloniert, exprimiert und aus *Escherichia coli* als heterologem Wirt aufgereinigt werden. Die beiden LOV-Proteine PpSB1-LOV und PpSB2-LOV binden oxidiertes Flavinmononukleotid als Chromophor und zeigen dieselbe primäre Photochemie wie die pflanzlichen Phototropin-LOV-Domänen. Interessant ist, dass die auf Aminosäure Ebene zu 66% identischen Proteine vollkommen verschiedene Dunkelrückkehr-Kinetiken aufweisen. Bei PpSB1-LOV handelt es sich, mit einer Rückkehr-Zeitkonstante von circa 30 Stunden bei einer Temperatur von 20°C, um das langsamste bisher beschriebene bakterielle LOV-Protein. Im Gegensatz dazu besitzt PpSB2-LOV, mit einer Rückkehr-Zeitkonstante von circa 100 Sekunden bei 20°C, die schnellste bisher beschriebene Rückkehr-Kinetik unter den charakterisierten bakteriellen LOV-Proteinen. Beide Proteine verfügen zusätzlich zu der hoch konservierten LOV-Domäne lediglich über kurze N- und C-terminale Verlängerungen, tragen jedoch keine fusionierte Effektor-Domäne. Eine bioinformatische Analyse der LOV-Protein kodierenden Genom-Regionen zeigte, dass die entsprechenden LOV-Gene im *P. putida* Genom in direkter Nachbarschaft zu einer ganzen Reihe von wahrscheinlich unter Eisenmangelbedingungen benötigten Genen lokalisiert sind. Da in Bakterien oftmals funktionell voneinander abhängige Gene gebündelt in Gen-Clustern vorliegen, wurde eine eventuelle Beteiligung der Blaulicht-Rezeptoren an der Aufnahme von Eisen untersucht. Tatsächlich zeigten erste photophysiologische Untersuchungen, dass unter Blaulichtbeleuchtung bei gleichzeitigem Eisenmangel, im Vergleich zu einer in Dunkelheit angezogenen Kultur, der Eisen-Siderophor Pyoverdin in größeren Mengen in den Kulturüberstand abgegeben wird. Bei Beleuchtung mit grünem und rotem Licht konnten

keine vergleichbaren Effekte beobachtet werden. Dies legt die Beteiligung (mindestens eines) Blaulicht-Rezeptors nahe.

Zusätzlich wurden innerhalb dieser Arbeit vergleichende Sekundär-(mittels Zirkulardichroismus-Spektroskopie) und Quartärstrukturanalysen (mittels Größenausschluss-Chromoatographie) an verschiedenen bakteriellen LOV-Proteinen durchgeführt. Die Ergebnisse dieser Untersuchungen unterstreichen die Bedeutung der LOV-Dimerbildung für die LOV-LOV sowie für die LOV-Effektor Interaktion. Außerdem wurde mit Hilfe dieser Analysen sowie mittels computergestützter Docking-Simulationen die Oberfläche, welche sich aus den fünf anti-parallelen verlaufenden β-Faltblättern der LOV-Domäne zusammensetzt, als Interaktions-Hotspot identifiziert. Für einige eukaryotische LOV-Domänen konnte durch strukturelle Untersuchungen auf tertiär- und quartärstruktureller Ebene [6-8] ebenfalls eine Dimerbildung gezeigt werden. Demzufolge legen diese Analysen nahe, dass in den eukaryotischen Systemen die Interaktion zwischen den entsprechenden LOV-Monomeren wahrscheinlich über dieselbe Proteinoberfläche wie in den prokaryotischen Systemen erfolgt. Damit könnte dieses Interface für den Signalweiterleitungsprozess sowohl in bakteriellen als auch eukaryotischen LOV-Proteinen von Bedeutung sein.

c) Biotechnologische Anwendung bakterieller LOV proteine

Im Zuge dieser Arbeit konnten zwei bereits charakterisierte bakterielle LOV-Protein Konstrukte (YtvA und dessen isolierte LOV-Domäne), sowie eins der neu identifizierten LOV-Proteine aus *P. putida* (PpSB2-LOV), mittels ortsgerichteter Mutagenese für die biotechnologische Anwendung als fluoreszierendes Reporter-Protein optimiert werden. Dabei wurde in dem kanonischen LOV-Sequenzmotiv GXN**C**RFLQG der photoaktive Cysteinrest (fett hervorgehoben) durch einen nicht-polaren Alaninrest ersetzt. Dies führt zur Unterbrechung des Photozyklus und damit, bei Anregung mit Blaulicht, zu einem kontinuierlichen Übergang des Flavin-Kofaktors zwischen dem Grundzustand und dem ersten angeregtem Singulett-Zustand. Die Rückkehr des Flavin-Kofaktors aus dem Singulett-Zustand in den Grundzustand erfolgt dabei über radiative und nicht-radiative Prozesse, wobei Photonen in Form von Fluoreszenz emittiert werden. In LOV-Proteinen ist die Fluoreszenz-Emission, im Gegensatz zum Grün Fluoreszierenden Protein (GFP), in welchem die Synthese des fluoreszierenden Chromophores nur unter Umsatz von Sauerstoff stattfinden kann, unabhängig von der Verfügbarkeit von Sauerstoff. Deshalb bieten sich die im Rahmen dieser Arbeit entwickelten, auf bakteriellen LOV-Proteinen basierenden, Fluoreszenz-Reporter- Proteine erstmals für die nicht invasive Visualisierung von anaeroben biologischen Systemen an. Solche Systeme waren bisher mittels GFP und dessen Varianten nicht oder nur schwierig zu analysieren.

List of Publications

Krauss, U., A. Losi, W. Gärtner, K.E. Jaeger, and T. Eggert. (2005). Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct. Phys Chem Chem Phys **7:**2804-11.

Krauss, U., and T. Eggert. (2005). insilico.mutagenesis: a primer selection tool designed for sequence scanning applications used in directed evolution experiments. Biotechniques **39:**679-82.

Buttani, V., A. Losi, T. Eggert, **U. Krauss**, K.E. Jaeger, Z. Cao, and W. Gärtner. (2007). Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions. Photochem Photobiol Sci **6:**41-9.

Drepper, T., T. Eggert, F. Circolone, A. Heck, **U. Krauss,** J.K. Guterl, M. Wendorff, A. Losi, W. Gärtner, and K.E. Jaeger. (2007). Reporter proteins for in vivo fluorescence without oxygen. Nat Biotechnol **25:**443-5.

Krauss, U., B.Q. Minh, A. Losi, W. Gärtner, T. Eggert, A.v.H. Haeseler and K.E. Jaeger. (2007). LOV is all around - A phylogenetic study on the origin of LOV-dependent blue-light signaling proteins. *Submitted for Publication*

Patent applications

Eggert, T., Drepper, T., Guterl, J.-K., Heck, A., **Krauss, U.** and Jaeger, K.-E. (2005) Fluoreszenzmarker und dessen Verwendung. Patent Application DE 10 2005 048 828.5

Abbreviations

In the following the non-standard abbreviations that were repeatedly used throughout this thesis are summarized. Excluded from this list are standard English-language abbreviations and SI units. Amino acids were abbreviated using the common one- or three letter codes.

aa	Amino acid
Acphot	Adiantum capillus-veneris phototropin
ADO	ADAGIO
Asphot	Avena sativa phototropin
ATP	Adenosine 5’-triphosphate
Atphot	Arabidopsis thaliana phototropin
BFP	Blue Fluorescent Protein
BLUF	Sensor of Blue-Light Using FAD domain
CD	Circular dichroism
CFP	Cyan Fluorescent Protein
Crphot	Chlamydomonas reinhardtii phototropin
CRY	Cryptochrome
C-terminal	Carboxy-terminal
DNA	Deoxyribonucleic acid
DsRed	Discosoma species Red (fluorescent Protein)
FAD	Flavin Adenine Dinucleoctide (Riboflavin 5'-adenosine diphosphate)
FbFP	FMN based Fluorescent Protein
FKF1	Flavin-Binding Kelch Repeat F-Box Protein
FMN	Flavin Mononucleotide (Riboflavin 5′-monophosphate)
GFP	Green Fluorescent Protein
GTP	Guanosine 5’-triphosphate
HK	Histidine Kinase domain
HPLC	High Performance Liquid Chromatography
HTH	Helix- Turn- Helix DNA binding domain
LOV	Light, Oxygen, Voltage domain
LOV390	LOV signaling state
LOV447	LOV dark state
LOV660	LOV excited triplet-state
LOV-HK	LOV Histidine Kinase
LOV-HK-RR	Hybrid LOV Histidine Kinase
N-cap	Amino-terminal cap
NMR	Nuclear Magnetic Resonance
N-terminal	Amino-terminal
PAC	Photoactivated Adenylyl-Cyclase
PAS	Per, Arndt, Sim domain
Phot	Phototropin
PPIX	Protophorpyhrin IX
PVD	Pyoverdine
PYP	Photoactive Yellow Protein
RF	Riboflavin
ROS	Reactive Oxygen Species
RR	Response Regulator
STAS	Sulfate Transporter Anti-Sigma factor antagonist domain
TG	Transient Grating
TrL	Transient Lensing
UV	Ultra-Violet
WC-1	White-Collar-1
YFP	Yellow Fluorescent Protein
ZTL	ZEITLUPE

1 Introduction

One theory of old says:

"And God said, "Let there be light"; and there was light. And God saw that the light was good; and God separated the light from the darkness." [1]

Some 14 billion years ago - According to the more recent theory …

" … of the Big Bang, the Universe started hot and dense and then expanded and cooled. In the hot, dense conditions of the early Universe, photons were tightly glued to matter. When the Universe was about 300,000 years old the temperature dropped below 3,000 K, allowing atomic hydrogen to form and releasing the photons. These photons, which traveled freely through the Universe as it expanded and cooled, make up the cosmic microwave background (CMB) we see today. Ten to twenty billion years after the Big Bang, the CMB is a cold sea of photons with an average temperature of 2.7 K (-270 °C). These photons are all around us, causing about 1% of the noise on our television sets." [2]

Some 9 billion years later:

"A star is born by the gravitational collapse of a cloud of dust and gas. If the condensing spherical mass cannot grow large enough, then it slowly gets colder and fades, becoming a brown dwarf. If it grows enough to sustain hydrogen fusion, a main sequence star, like our sun, is formed." [3]

In what way light or in a more principal way photons came into existence during the early days of our universe is still far from resolved and a matter of personal belief and faith. However, one thing is unquestionable. The small part of the overall radiation that is emitted by our sun and that is called visible light, extending from 380 to 780 nanometers, has been crucial for the evolution and survival of life on earth as we know it today. All forms of life on earth are in one way or the other dependent on radiation energy for their survival and maintenance. Certain bacteria, algae and higher plants utilize this radiation energy directly for the synthesis of essential

[1] Moses, *Genesis, Chapter 1: The Creation*, in *The Holy Bible, containing the Old and New Testaments, King James Version*. 1999, American Bible Society: New York

[2] Hu, W., *Ringing in the new cosmology.* Nature, 2000. **404**: p. 939-40
[2] Rowan, L., *Stellar Birth and Death.* Science, 1997, **276**: p. 1315

components by employing the photosynthetic process. We, as all animals, cannot utilize this radiation energy directly but depend on the consumption of plant materials or other animals that consumed material derived from plants for their survival. Thus, the ultimate source of energy on our planet is the sun-(light) and photosynthesis is essential for maintaining many forms of life on our planet [9]. It is therefore not surprising that all photosynthetic organisms have developed signaling systems to sense their source of energy. Such photo- or light receptors enable them to optimally utilize the process of photosynthesis for energy generation by not only sensing the direction of the light source but also integrating its spectral properties, intensities and the duration of light availability as a signal. In particular the spectral properties of light that living organisms encounter, depending on their habitat, are variable. Hence, it is not surprising that several different photoreceptors have emerged or were "invented" by evolution to facilitate the adaptation to such changing light conditions.

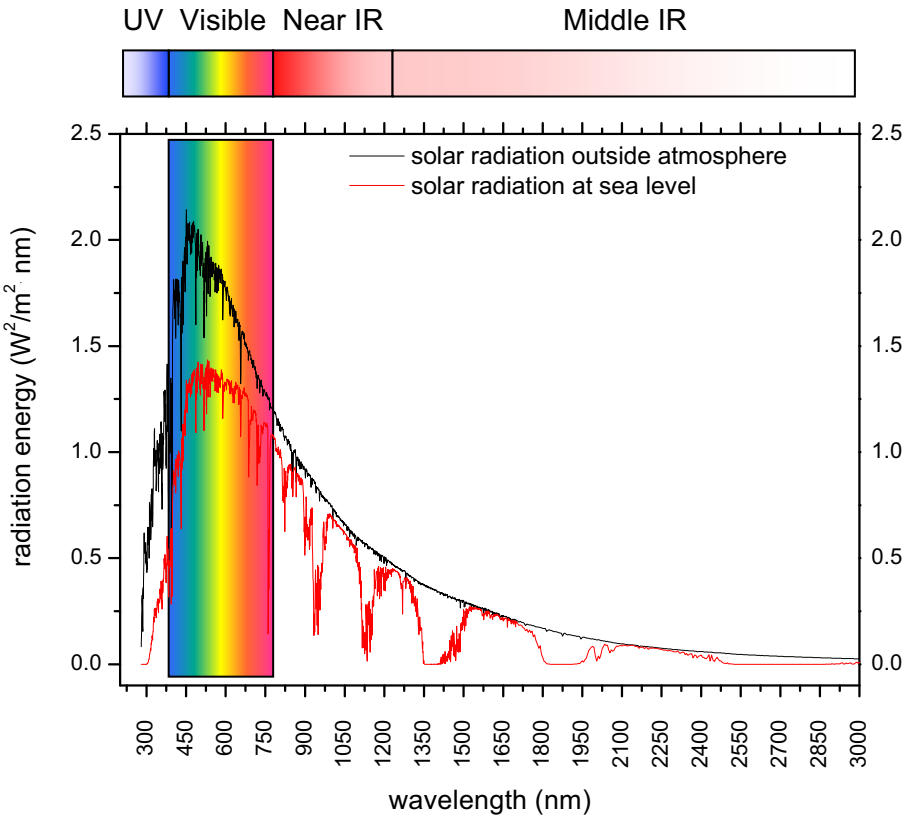


Figure 1: Reference solar spectra of the American Society for Testing and Materials (ASTM) as implemented in the standard ASTM G-173-03 [10]. The recorded sea level spectrum is shown in red. The solar spectrum outside of the atmosphere (in black) was modeled by taking into consideration the 1976 US standard atmosphere [11].

The radiation with the highest inherent energy that reaches the surface of the earth, and therefore most strongly influences all living beings, is the visible fraction

(highlighted in Figure 1). Thus, all photoreceptor proteins so far described in the three kingdoms of life allow responses to very distinct wavelengths in the visible spectral region.

1.1 The classification, distribution and physiological role of photoreceptors and the conserved photosensing paradigms

Light, or more principally photons, consist of energy in the form of oscillating electric- and magnetic fields. Likewise electrons and protons, that ultimately constitute molecules and matter, are charged particles and hence their motions generate oscillating electric fields. A material can only absorb energy from light (photon-capture) if the frequency of the light oscillation and the vibration frequency of the electrons in the material match. To enable photoreceptor proteins to respond to visible radiation, light absorption has to occur in the protein.

The ability of photoreceptor proteins to absorb visible light is, at the molecular level, directly correlated to the type of chromophore (from the Greek word *chromos* = color) that is bound within the respective photoreceptor protein.

The feature that characterizes a given chromophore is its ability to absorb photons of the incident visible radiation, resulting in the transition of electrons in the chromophore to a higher energy level and eventually loss of the photon and attenuation of the light. The transition of electrons in the chromophore and thus photon capture can only occur when the energy difference between two molecular orbitals of the chromophore matches the photon energy of the incoming radiation. Macroscopically, the intensity of light absorption by atoms and molecules depends on the transition probability of the electrons in the chromophore, which is phenomenologically described by its absorption coefficient.

Therefore, the biologically most useful criteria for the classification of photoreceptor proteins into families is to divide them according to the wavelength-fraction which the chromophore of the photoreceptor protein absorbs maximally. Correspondingly, this absorption maximum in turn determines the fraction of the visible radiation the protein optimally responds to (Figure 2).

The corresponding biological response, initially trigged by a given photoreceptor protein, is thus initiated by a chain of events: photon-capture in the chromophore leading to the primary photochemical event, e.g. photoisomerization of a double bond

in retinal [12]. This in turn results in conformational changes in the photoreceptor protein that are propagated from the chromophore moiety to so-called output/effector domains that can be fused protein modules or protein-protein interaction partners of the photoreceptor. Those output or effector domains eventually trigger the biologically relevant events such as for example regulation of gene expression, protein degradation, antagonistic partner-switching in terms of protein-protein interactions etc.

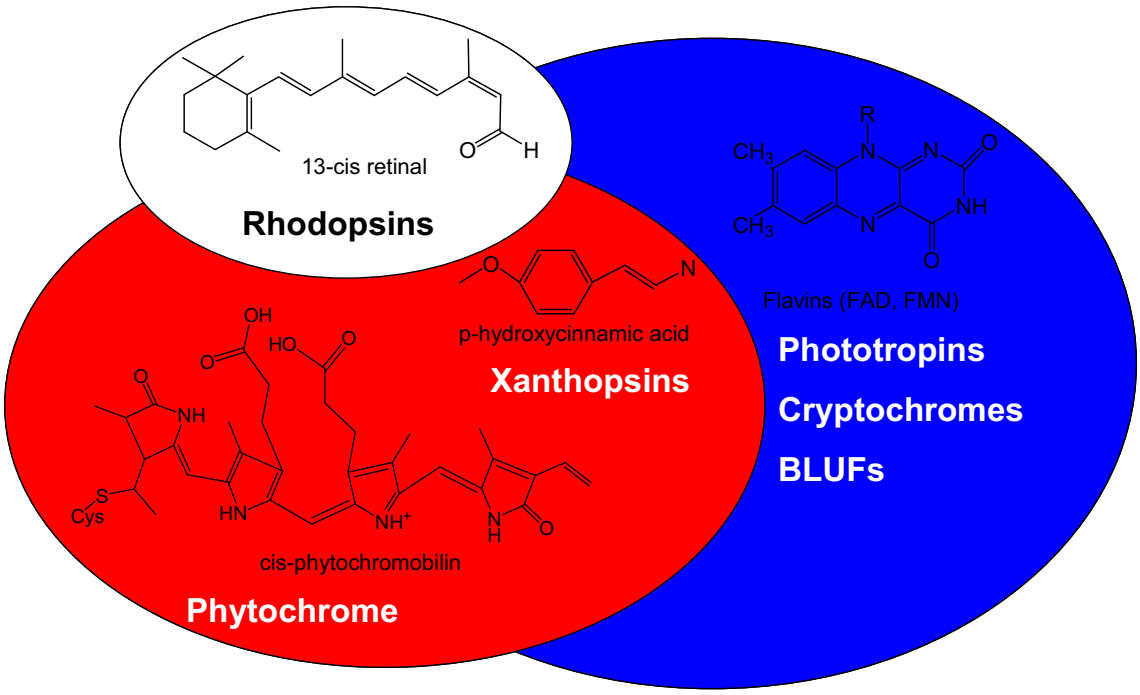


Figure 2: Photosensing paradigms. Classification of photoreceptor families according to the light sensitive chromophore that is utilized for photon-capture; phototropins, ZTL/ADO/FKF1, cryptochromes and the BLUF photoreceptors all bind similar flavin co-factors (flavin adenine dinucleotide (FAD) or flavin mononucleotide (FMN); for a detailed structural representation of the respective flavin molecules see Appendix A) but undergo very distinct photochemical reactions, whereas the other photoreceptor families, the phytochromes, the xanthopsins and the rhodopsins bind very different co-factors that basically undergo the same photochemical reaction upon photon-capture, namely a *cis/trans* isomerization of a double bond of the respective co-factors. For the phytochrome family, only the phytochromobilin co-factor of the plant phytochromes-sub-familiy is shown for the reason of simplicity.

Currently, seven photoreceptor families are distinguished and are sometimes found widely distributed in the three kingdoms of life (Figure 3): rhodopsins[13], phytochromes [14], xanthopsins [15], phototropins [16], ZEITLUPE(ZTL)/ ADAGIO(ADO)/Flavin-binding Kelch-Repeat F-Box protein(FKF1) [17], cryptochromes [18], and BLUFs (sensor of Blue-Light using FAD) [19]. In the first three families, photon absorption causes a *cis/trans* isomerization of a double bond of the three different chromophores. The latter four families use riboflavin derivatives as chromophore, (with different modifications/substitutions) but each family

undergoes distinct photochemial reactions, whereas the phototropin and the ZTL/ADO/FKF1 family both contain so-called light, oxygen, voltage (LOV) domains as the light-sensitive sensor module [20].

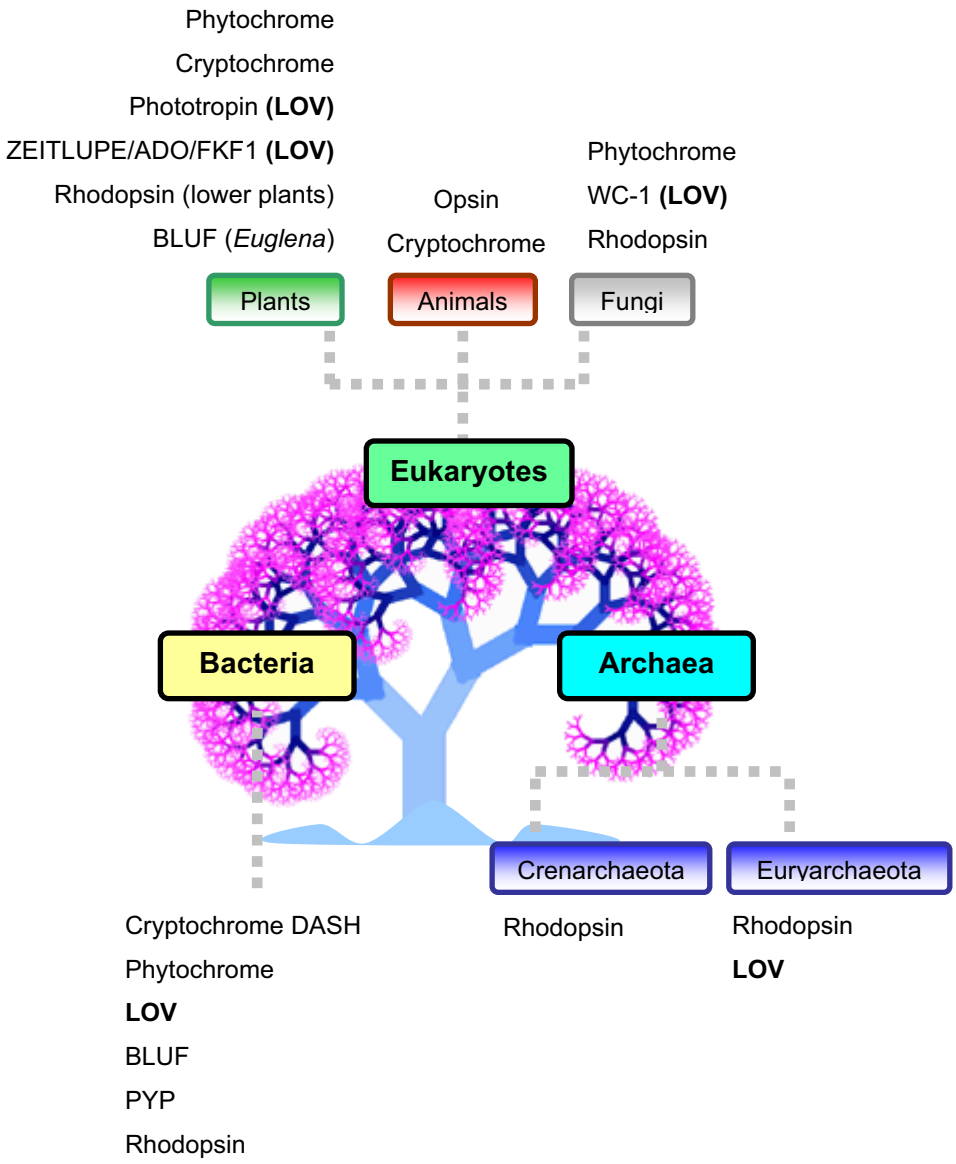


Figure 3: Distribution of photoreceptor proteins throughout the three domains of life. Photoreceptor proteins of the six mentioned families are widely distributed throughout all kingdoms of life and hence can be found within organisms of a variety of habitats. The LOV domain-containing photoreceptor proteins (e.g. plant phototropins) can be found in Archaea, Bacteria and Eukaryotes (with the exception of animals). The other photosensory proteins e.g. cryptochromes, phytochromes and rhodopsins are equally widely distributed. Abbreviations: LOV: light, oxygen, voltage domain family of photoreceptors (include plant phototropins and ZEITLUPE(ZTL)/ADAGIO(ADO)/Flavin Binding Kelch Repeat F-Box proteins(FKF1), BLUF: Sensor of blue-light using FAD domain, PYP: photoactive yellow protein, Cryptochrome DASH: family of cryptochromes, whereas the name underscores the relationship with cryptochromes found in *Drosophila*, *Arabidopsis*, *Synechocystis*, and *Homo* (DASH)

Classical rhodopsins, binding an isomerizable retinal chromophore, are found in Archaea, e.g. in the halophilic *Halobacterium salinarum* [21-23]. Up to now no rhodopsin homolog was found in higher plants but it could be shown that the unicellular green alga *Chlamydomonas reinhardtii* possesses two rhodopsins mediating tactic responses to high- and low light intensities [24]. Furthermore, rhodopsins are found as rod- and cone visual pigments in the retina of animals [25] or participate as so called opsins (e.g. melanopsin) in setting the mammalian circadian clock [26, 27]. Due to the increasing number of completely sequenced fungal and microbial genomes more and more putative retinal binding rhodopsins were discovered in several fungi, in some γ-proteobacteria as well as in the α-proteobacterium *Magnetospirillum magnetotacticum* [28] and in the cyanobacterium *Anabaena (Nostoc)* sp. PCC7120 [29].

The red/far-red sensing, tetrapyrrole ligand-binding phytochromes are widely distributed throughout the bacterial and eukaryotic domains of life but are apparently absent in the archaeal kingdom. Their presence could be verified in higher plants (phytochromobilin binding), in cyanobacteria (phycocyanobilin binding), as well as in some proteobacteria (predominately plant pathogens or symbiotic living microbial species). Recently, they were also discovered in the fungal kingdom [30, 31]. The latter two classes bind the tetrapyrrole biliverdin via a thioether-bond to a conserved cysteine residue in a PAS (Per, Arndt, Sim) domain whereas the plant and cyanobacterial phytochromes bind the tetrapyrrole ligand via a conserved cysteine residue in the phytochrome GAF (cGMP phosphodiesterase/adenylate cyclase/ FhlA) domain [32]. In plants phytochromes control cellular responses and tropisms such as chloroplast movement, cytoplasmic motility, endoreduplication (genome duplication without mitosis), and nyctinastic movements of leaves (opening and closing of leaves with a circadian rhythm), as well as other tropic responses such as gravitropism, polarotropism, and phototropism [33]. In contrast to the physiologically well characterized plant phytochromes, their prokaryotic counterparts from cyanobacterial and other prokaryotic genera remain in most cases unlinked to red/ far-red light dependent photomorphogenic responses *in vivo* [34]. However, some phytochromes and phytochrome-related proteins from prokaryotic genera have been linked to light dependent physiological responses *in vivo*. For example, the first prokaryotic phytochrome-like protein RcaE that was identified in the cyanobacterium *Fremyella diplosiphon* (also referred to as *Calothrix* sp. PCC 7601), controls the process of

complementary chromatic adaptation (CCA) [35]. A homolog of RcaE, PlpA (phytochrome-like protein) from *Synechocystis* sp. PCC 6803 is required for growth of *Synechocystis* under blue light [36].
Interestingly, the *trans p*-coumaric acid binding xanthopsins [15], whose archetype was the photoactive yellow protein (PYP) isolated from *Ectothiorhodospira halophila* [37], seem to be much less widespread than the other photoreceptor proteins. Its presence seems to be restricted to proteobacterial genera. Furthermore, no conclusive role(s) could be assigned to those otherwise photochemically and structurally well characterized photoreceptor modules [38]. It was proposed that *E. halophila* PYP functions to regulate negative phototaxis away from harmful UV light [39], but the regulatory role of PYP in this process could not be proven until today.

The fraction of the visible radiation with the highest inherent energy (excluding the near-UV and far-UV light) is the blue region of the spectrum (Figure 1) (430-500nm, energy content approximately 239 to 280 kj/mol), hence the light with the most profound impact on all living beings is probably blue-light. Another feature that distinguishes the blue-region from the other wavelengths of the visible spectrum and hence contributes largely to its impact on our planet, is its high penetrability into the oceanic water column [40]. For example, open ocean deep waters (100-200m depth) are enriched with blue-light, whereas the longer wavelengths of light, e.g. red-light, do not penetrate below approximately 15m [40]. Organism responses to blue-light might either facilitate the optimal utilization of the photosynthesis process for energy generation (photosynthetic organisms) or on the other hand initiate tactic light avoidance responses (photophobie) to protect the organisms from the harmful effects of energy-rich blue-light or UV-light. UV-light is well known as a causative agent of severe DNA-damage in living specimens [41]. Contrary, a harmful effect of blue-light might be mediated by its capability to excite, with high yield, ubiquitously present photosensitizing compounds, e.g. porphyrins and flavins [42, 43]. Photoexcitation converts such compounds with high efficiency into the triplet state which in the presence of oxygen generates the highly oxidative oxygen singlet state and other reactive oxygen containing species (ROS) which in turn can cause damage to a wide range of cellular tissues and organs [42].
Therefore, it is not surprising to find several photoreceptor families that respond to this region of the visible radiation.

The flavin-containing blue-light photoreceptors of the cryptochrome family are found in lower and higher eukaryotes [including mammals (*Homo sapiens*), insects (*Drosophila*), plants (*Arabidopsis*), and algae (*Chlamydomonas*)] [18, 44]. The classical plant or animal cryptochromes are characterized by their high degree of sequence similarity to DNA-photolyases, but lack intrinsic DNA-photolyase activity [45-47]. DNA photolyases are enzymes that utilize blue-light to repair UV-induced DNA damage by removing pyrimidine dimers from double stranded (ds) DNA [47]. In recent years, a new class of putative cryptochromes was discovered in several photosynthetic and non-photosynthetic prokaryotes, e.g. in *Synechocystis* sp. PCC6803, [48, 49] and in *Vibrio cholerae* [50]. This new type of cryptochrome familiy was referred to as CRY-DASH, to highlight its relationship with cryptochromes found in *Drosophila*, *Arabidopsis*, *Synechocystis*, and *Homo* (although CRY-DASH itself is not found in *Drosophila* or humans) [51]. With the exception of the CRY-DASH familiy, to which no blue-light dependent regulatory role could so far be assigned, the cryptochromes are involved in processes ranging from synchronization of the circadian clock in animals to hypocotyl elongation, seed germination, and pigment accumulation in plants [46, 52]. For members of the CRY-DASH family, it was recently demonstrated that they possess DNA- photolyase activity towards cyclobutane pyrimidine dimers in single stranded (ss)-DNA, but lack significant repair activity on ds-DNA [53]. This discovery challenges the placement of the CRY-DASH family of proteins among the classical cryptochrome blue-light photoreceptors.

A recent discovery was the Blue-Light sensing Using FAD (BLUF) family of photoreceptor proteins. BLUF seems to be predominant among prokaryotic (proteobacterial and a few cyanobacterial) genera. Among the eukaryotes, homologous proteins could so far only be found in the unicellular flagellate *Euglena gracilis* and other Euglenoids [19, 48]. The BLUF-protein prototype, AppA from *Rhodobacter sphaeroides*, is involved in blue-light dependent regulation of the photosynthesis genes in this phototrophic purple bacterium [54]. In *Euglena*, the BLUF domain-containing photoactivated adenylyl cyclase (PAC) controls both negative and positive phototaxis as well as photophobic (abrupt turn in response to a rapid increase [step-up] or decrease [step-down] in the light fluence rate) responses [55, 56].

1.2 The light sensitive LOV signaling modules in plant photoreceptors

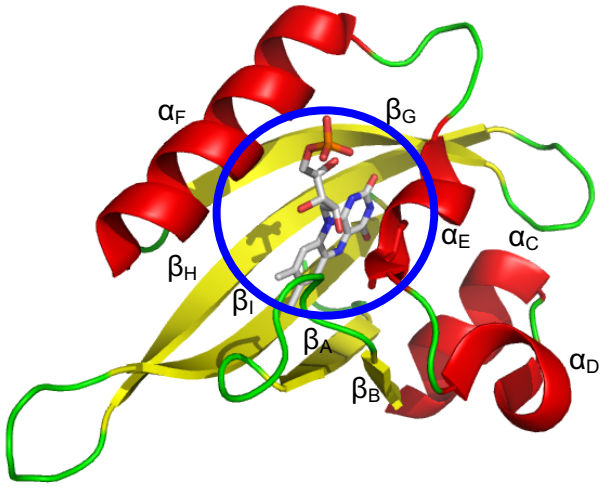
Already Charles Darwin and his son Francis described a blue-light dependent hypocotyl movement for several plant species including, *Arabidposis* and *Zea mays,* in their classical volume "The power of movement in plants", published in 1880 [57]. Nevertheless, it took more than one century until Winslow Briggs and colleagues were finally able to trace Darwin's observation to its molecular origin. They identified a gene named *nph1,* whose gene product was acting very early in the signal-transduction chain for phototropism in *Arabidopsis thaliana* [16]. The corresponding gene product was later named phototropin, after the response it initiates [20]. After the identification of the phototropins (phot1 and phot2) as the primary blue-light photoreceptors for plant phototropism [58], it soon became clear that they control other phenomena such as chloroplast movement, leaf expansion, and stomatal opening in a blue-light dependent manner [59]. In the green alga *Chlamydomonas reinhardtii*, phot is instead involved in the control of multiple steps in the sexual life cycle [60]. Later biochemical and photochemical studies demonstrated that phototropin contains two flavin-binding LOV domains (designated LOV1 and LOV2, respectively) which function as blue-light sensitive signaling modules that cause light dependent autophosphorylation of a phot-coupled serine/threonine kinase [1], and hence enable downstream signaling, probably via a phospho-relay to yet unidentified regulatory proteins.

Another class of LOV domain-containing photoreceptor proteins that is found predominantly in higher plants is the ZTL/ADO/FKF1-family. This family of LOV photoreceptor proteins plays a primary role in the photocontrol of flowering time [61] and the circadian period in higher plants [17].

1.3 The LOV paradigm – conserved structure and photochemistry

The blue-light sensitive LOV domains of phototropins share with other signaling modules, e.g. with the light-sensitive photoactive yellow protein (PYP) that binds an isomerizable *p*-coumaric acid chromophore, the heme-binding, oxygen-sensing FixL protein as well as with the FAD-binding redox-sensing NifL protein, a common structural folding motif, the so-called α/β PAS fold. The LOV domain exhibits an α/β fold with the following arrangement of the secondary structure elements: βAβB-

αCαDαEαF-βGβHβI (Figure 4). The central anti-parallel β-scaffold is constituted by two distinct portions βAβB and βGβHβI, which are linked by a helical connector, αCαDαEαF, together forming a pocket in which the chromophore flavin mononucleotide (FMN) is non-covalently anchored in the dark state. The helical connector that comprises the canonical sequence motif (GXNCRFLQ), harbors the photoactive cysteine residue (highlighted in bold), responsible and indispensable for the LOV photoreaction mechanism.




```
                  *         20        *         40        *         60        *         80        *        100       *
VIVID LOV : CALIICDLKQKTPIVYASTAFLYMTGYSNAEVLGRNCRFLQSPDGMVKPKSTRKYVDSNTINMRKAIDRNA-EVQVEVVNFKKNGQRFVNFLTMIPVRDETGEYRYSMGFQC :
CreinLOV1 : HTFVVADATLPDCPLVYASEGFYAMTGYGPDEVLGHNCRFLQGEG----------TDPKEVQKIRDAIKKCE-ACSVRLLNYRKDGTPFWNLLTVTPIKTPDGRVSKFVGVQV :
Phy3LOV   : NSFIVVDALKPDFPIIYASTGFLNLTGYTSREVIGQNCRFLQGPD----------TNPADVASIRFALAQGTGTFCGRLLNYRKDGSSFWNLLTIAPIKDLGSIVKLIGVQL :

                  βA        βB      αC    αD      αE     Insertion          αF          βG          βH              βI
Phy3-LOV2   cEEEEEEccc-cccEEEEEEcHHHHHHcccHHHHcccccHHcccc----------ccccHHHHHHHHHHHHcccEEEEEEEEE-cccEEEEEEEEEEEEc-ccccccEEEEE
```

Figure 4: 3D-Ribbon representation of the LOV domain fold. Here, the dark state of *C. reinhardtii* LOV1 (pdb entry: 1N9O) is shown, with alpha-helices depicted in red and beta-strands in yellow. The assignment of the secondary structure elements is also shown. Below the structure a sequence alignment of the three LOV domains for which a structure is available is shown: the LOV domain of the *N. crassa* VIVID protein (2PD7), *C. reinhardtii* LOV1(1N9O) and the LOV domain of the chimeric phytochrome-phototropin fusion (Phy3, 1G28) of the fern *Adiatum capillus veneris*. Very recently the crystal structure solved for the LOV domain of the bacterial photoreceptor YtvA (2PR5) unequivocally proved the conservation of the LOV fold, also among bacterial genera.

The LOV-mediated light-signaling mechanism (or the LOV paradigm; Figure 5) is highly conserved among all currently photochemically characterized LOV proteins. In brief, photon absorption in the FMN molecule that is bound non-covalently in LOV domains dark state (LOV447, numbers refer to the absorption maximum of the particular intermediate), excites the FMN molecule on a ps- time scale [62] to its excited triplet-state ^{3}[FMN](LOV660). The signaling state (LOV390) formation occurs via the decay of LOV660, typically within 1-2μs [63, 64]. This signaling state formation is reached by the formation of a covalent bond between the position 4a of

the flavin chromophore and the SH-group of the essential cysteine in the above mentioned conserved sequence motif. The longest living species of the LOV photocycle, the signaling state intermediate (LOV390), thermally recovers in the dark to the ground state within minutes to hours for plant and bacterial LOV domains [3, 65-68]. This recovery can even require several days as reported for the LOV domain of the Flavin-Binding-Kelch-Repeat F-box protein of *Arabidopsis* [67]. For a more detailed review of the primary events in the LOV photocycle, as well as for a summary of current mechanistic proposals, see the comprehensive review by Losi (2007) [69].

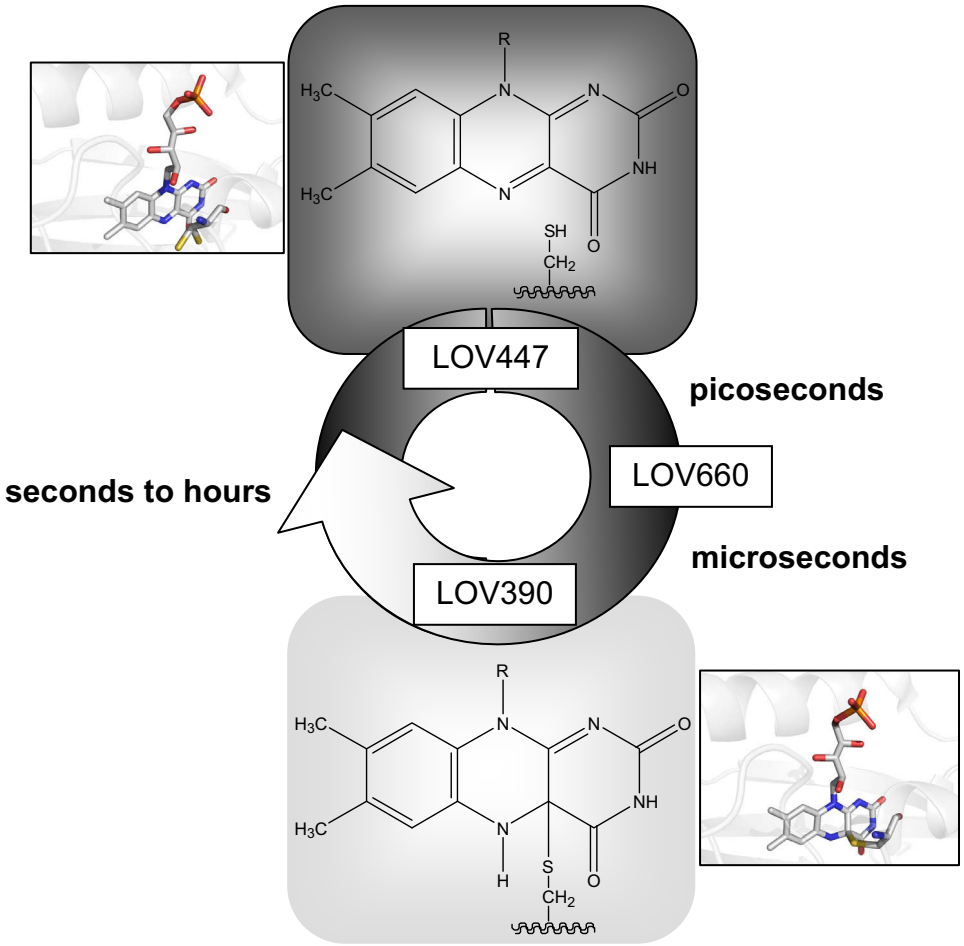


Figure 5: Simplified scheme illustrating the LOV photocycle. The dark state (LOV447) after blue-light illumination gives rise within picoseconds to a red-shifted triplet-state intermediate (LOV660). The LOV signaling state (LOV390) is formed by a decay of the transient triplet-state intermediate on a microseconds timescale. Dark state recovery occurs thermally within seconds, hours or even days.

Although much is already known about the early events in the LOV photocycle, leading from photon absorption in the FMN molecule to the signaling state formation in the protein, quite as much remains unknown about the subsequent processes finally leading to the physiological output. A number of open questions can be listed such as: i) which initial conformational change(s) in the LOV domain finally triggers

the biological output?, ii) how is the signal, generated in the LOV core, relayed to the corresponding output domain(s) (i.e. to the kinase in plant phot)?, iii) how does the dark recovery proceed mechanistically and why are the recovery kinetics so different for various LOV domains, especially among the bacterial LOV systems? Regarding these pressing questions, several mechanisms were suggested during the last year, which will be discussed in detail in the following chapter.

1.4 The proposed signal-transduction mechanisms in plant phot-LOVs

Currently the lack of any structural information for any full-length LOV domain-containing protein that includes - apart from the conserved LOV core - a fused effector domain (e.g. phototropin consisting of LOV1, LOV2, and the serine/threonine kinase) clearly hampers the study of the structural basis of the signal relay from the light sensitive LOV domain to coupled effector domains.

In addition, X-ray structures determined for the dark and light state of single LOV domain modules such as LOV2 of the fern *Adiantum capillus-veneris* and LOV1 of the green-alga *Chlamydomonas reinhardtii* did not yield much information regarding the downstream signaling mechanism as both dark and light state crystal structures did not differ dramatically, whereas the only major changes were restricted to the region surrounding the flavin chromophore. This led to the suggestion that signal-transduction mechanisms must be dynamic in nature [70]. Nevertheless, several hypotheses were brought forward regarding how the light signal that is received in the LOV domain might be transmitted:

i) A conserved salt-bridge between E51 (on αD) and K92 (on the βG-βH loop) (Crphot-LOV1 numbering) was suggested to be involved, whose stability may modulate changes in the binding affinity between the LOV domain and its partner domains [71]. This hypothesis was recently challenged through Molecular Dynamic (MD) simulations for the LOV1 and LOV2 domains of *C. reinhardtii* phot as well as through several mutational studies on different phototropins (as elaborated in the following paragraph).

In the case of LOV1, MD simulations indicated that the E-K salt bridge is broken in the dark state but formed in the light state, whereas in LOV2 the E-K salt bridge does not undergo major changes between dark and light states [72]. Hence, the authors suggested a functional role for the E-K salt-bridge in the photoreactivity of LOV1 but

not for LOV2, respectively. However, regarding phot function, only LOV2 photoreactivity seems to be essential whereas LOV1 photoreactivity is largely dispensable [73]. Together with the aforementioned suggestion that the E-K salt-bridge might not undergo changes between light and dark states of LOV2 observable in an MD simulation, several lines of evidence now suggest that the phot activation does not occur via a mechanism involving the conserved E-K salt-bridge. A mutational study further supports the latter argument experimentally. In this study, the authors could show that the disruption of the E-K salt in the LOV2 domain of full-length *Arabidopsis thaliana* phot1 does not affect the light-induced activity of the protein [74]. Nevertheless, the E-K salt bridge could still be functionally important for phot signal-transduction by stabilizing the LOV2 core during the photocycle and/or mediating dimerization of full-length phot through the LOV1 domain.

ii) Very recently a glutamine-flipping mechanism was suggested to play a role in LOV2 mediated phot activation. Studies on the functional role of the flipping-glutamine Q1029 (numbering follows phy3 from *Adiatum capillus-veneris*) or Q575 (numbering according to phot1 of *Arabidopsis thaliana*) suggested a functional role for this residue in light-driven phot kinase activation, as a mutation of this residue to leucine attenuates the light-induced autophosphorylation reaction [74] and furthermore impairs light-driven conformational changes in the central β-sheet region [75].

iii) Harper and colleagues performed NMR experiments on an extended LOV2 construct of *A. sativa* phot1 that contained C-terminally to the LOV2 core an extension of about 20 aa, which adopts a helical conformation (hence termed Jα-helix) [76]. They could demonstrate that this helical segment interacts tightly with the LOV core in the dark state but becomes unfolded and probably dissociates from the core LOV domain upon illumination [76]. Furthermore, mutational disruption of the interaction between the Jα-helix and the LOV core resulted in constitutive activation of the phot kinase domain [77]. A similar mechanism for another secondary structural element outside of the canonical LOV core was recently suggested for fungal blue-light receptor VIVID [68].

In the recently solved crystal structure of the short LOV sensor VIVID from *Neurospora crassa* an N-terminal cap (N-cap) that is partially helical in structure is found making extensive contacts with the central β-sheet of the core LOV domain. Although, the light-driven changes in the crystal of VIVID were very small. The

protein in solution showed a dramatic increase in its hydrodynamic radius, which was attributed to an increased disordering of the helical portion in the N-cap of VIVID [68]. Thus, a suggestion appears plausible that either the dissociation of the Jα-helix (in phototropin) or correspondingly the disordering of the N-cap (VIVID), both probably resulting in the exposure of the central β-scaffold of the LOV core, might trigger phot activation or, as in case of VIVID, enable downstream signaling (e.g. via protein-protein interactions).

1.5 The distribution of prokaryotic phototropin-like photoreceptors

In recent years genome mining and an increasing number of biochemical and biophysical studies have pointed towards shared photosensing paradigms between such distant taxa as e.g. bacteria, fungi, animals and plants [19, 29, 45, 66, 71, 78-81]. Hence, it was not surprising to eventually find plant phototropin-like (or LOV domain-containing) protein modules also in (oxygenic) photosynthetic prokaryotes such as in cyanobacteria and in some (anoxygenic) phototrophic proteobacteria such as e.g. *Rhodobacter sphaeroides* and *Erythrobacter litoralis* [66].
More surprising was the nearly ubiquitous distribution of LOV domain-homologous sequences in non-photosynthetic proteobacteria [66], since those organisms can not directly benefit from sensing a (blue)-light source as photosynthetic prokaryotes and obviously plants.
Today, among non-photosynthetic prokaryotes, LOV domain-homologous sequences can be found in a variety of plant pathogen species such as *Pseudomonas syringae, Xanthomonas campestris,* in some plant root colonizing Pseudomonads, like *Pseudomonas putida* and *P. fluorescens*, in human/mammalian pathogens such as *Listeria* and *Brucella* species, but also among common soil and leaves bacteria like *Bacillus subtilis*. A summary of the distribution of LOV domain-containing proteins among the three domains of life as well as the attempt to understand their phylogeny and the inherent evolutionary processes that might have contributed to the ubiquitous distribution of the LOV signaling module in today's biosphere are part of this thesis and are further discussed in Chapter 2.1 and Chapter 5.1.

1.6 Prokaryotic LOV proteins – a chromophore module servicing multiple output domains

The first photochemically and biochemically characterized LOV domain homolog of prokaryotic origin was the *B. subtilis* YtvA protein [3]. Losi, Gärtner and co-workers could demonstrate that the YtvA protein and its isolated LOV domain bind oxidized FMN and basically undergo the same light-induced photochemistry as plant-phototropin LOV domains (LOV1 and LOV2) [3]. Whereas the LOV domains of plant phototropins are invariably found to be fused to serine / threonine kinases as output

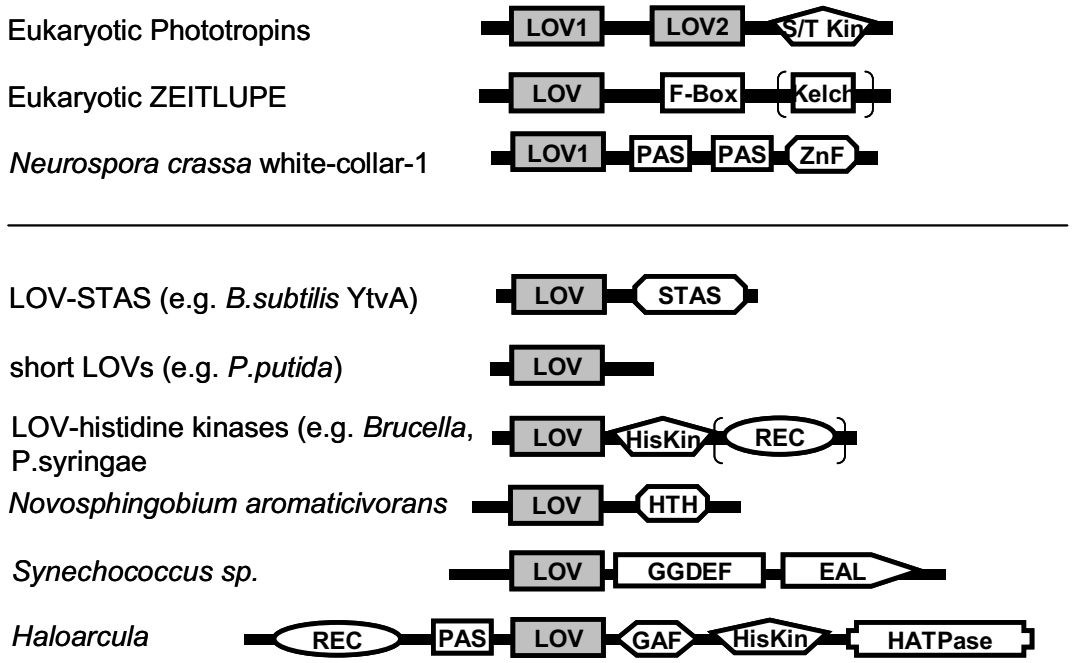


Figure 6: LOV domain-containing full-length protein architectures. The highly conserved eukaryotic photoreceptor families, namely the phototropins, the ZTL/ADO/FKF1 and the fungal white-collar-1 system of Neurospora crassa, are depicted in the upper half of the figure. Below the multitude of bacterial LOV domain-containing putative photoreceptor systems is shown.
Abrreviations: LOV: light, oxygen, voltage domain; S/T Kin: serine/threonine kinase; Kelch: Kelch repeats; PAS: Per, Arndt, Sim domain; ZnF: zinc-finger motif; STAS: sulphate-transporter antisigma-factor antagonist domain; HisKin: histidine kinase; REC: response regulator; HTH: helix-turn-helix DNA binding domain; GGDEF: diguanylate cyclase; EAL: phosphodiesterase; GAF: domain present in phytochromes and cGMP-specific phosphodiesterases; HATPase: histidine kinase ATPase domain.

domain, the bacterial LOV domain-containing protein architectures are much more variable with respect to their associated effector modules [66]. As an example, the first characterized prokaryotic LOV protein, YtvA of *B. subtilis*, possesses a sulfate-transporter anti-sigma factor antagonist (STAS) domain coupled to the conserved LOV core. This architecture is conserved among all the LOV domain-containing

proteins found in the sequenced Firmicutes genera (e.g. *Listeria monocytogenes* and *Oceanobacillus iheyensis*).

Another large group of LOV domain-containing proteins that possess a conserved full-protein architecture, comprising a LOV histidine kinase fusion, is constituted by sequences distributed in plant pathogen species (e.g. *Pseudomonas syringae* or *Xanthomonas* spp.), but are also found for example in the aquatic living *Caulobacter crescentus* and other α-proteobacterial genera, e.g. *Novosphingobium* and *Sphingomonas* species, as well as in human/animal pathogens like *Brucella*. That group of LOV proteins invariably carries a histidine kinase which is sometimes followed by an associated response regulator, fused to a single LOV domain. Previous studies on the *Caulobacter* LOV histidine kinase demonstrated a phot-like photochemistry for the isolated LOV domain, whereas the full protein could not be expressed in a soluble form [66]. Recent studies on another LOV histidine kinase from the plant pathogen *P. syringae* pathovar tomato showed, apart from the conserved phot-like photocycle, that the full-length protein undergoes, similarly to the plant phot-system, blue-light driven autophosphorylation in the (histidine) kinase domain, with subsequent relay of the phospho-group to the coupled response regulator. ([82], in press) and [83]. In the latter publication [83] three other LOV histidine kinases were described: from the human/animal pathogens *Brucella melitensis* and *Brucella abortus*, as well as from the marine phototroph

E. litoralis all showed a phot-like photochemistry as well as light driven autophosphorylation.

Other full-protein architectures that contain a LOV domain comprise for example diguanylate cyclase (GGDEF) [84, 85] and phosphodiesterase (EAL) domains. [85] Both classes of putative effector domains are reportedly involved in the turnover of cyclic-di-GMP, an emerging new class of global second messenger molecules [86]. Other full-protein organizations comprise helix-turn-helix (HTH)-transcriptional regulators [87], PAS domains [88] as well as GAF domains [89] that are usually found associated to phytochromes.

Some LOV protein architectures, mainly conserved among the saprotrophic fluorescent Pseudomonads such as *P. putida* and *P. fluorescens*, completely lack a fused output domain. Since those last mentioned systems are one main subject of the presented thesis, details are summarized in Chapters 3.1 and 3.3 and are furthermore discussed in Chapters 5.2.1 and 5.2.3

This impressive variety of putative effector domains, which can be found in the prokaryotic LOV proteins as well as different full-length architectures found in the different eukaryotic LOV photoreceptors (e.g. phototropin, fungal white-collar-1 (WC-1), and plant ZEITLUPE), raises the question whether the signal-transduction mechanism is conserved among all those highly different full-length LOV protein architectures. One might ask whether there is a common theme that governs the signal relay, e.g. a common interaction surface, or whether conserved dynamic conformational changes in the LOV domain exist that trigger the activation of different, structurally un-related effector domains? Some of this aspects are part of this thesis and are separately discussed in Chapters 3.2 and 5.2.2.

1.7 Biological role of prokaryotic LOV blue-light photoreceptors

Plant phototropins were originally identified using genetic screens for the loss of phototrophic hypocotyl elongation. Although the effect of phototropism was known already in Darwin's times, modern genetic tools finally facilitated the identification of the molecular basis of a long known phenomenon. In the case of prokaryotic photoreceptors of the LOV family, which are the subject of the presented thesis, the process was reciprocal. The proteins were primarily identified based on sequence similarity to plant phot-LOV domains. Subsequently, biophysical analysis proved their blue-light sensitivity and revealed the conservation of the LOV signaling paradigm throughout the three domains of life. However, studies regarding their biological significance were lacking behind the advances made to understand the LOV signaling mechanism. Only recently first hints emerged, demonstrating that some of those bacterial LOV domain-containing proteins (in particular the YtvA protein of *B. subtilis* [90-92] and a LOV histidine kinase from *B. abortus* [83]) mediate blue-light dependent physiological responses *in vivo*. Nevertheless, the available information is still scarce compared with the plant phototropins and the plant and fungal circadian LOV-photoreceptors (ZEITLUPE [17, 93] and white-collar-1 [94, 95]) for which detailed physiological responses as well as signaling networks are described.

1.7.1 The YtvA protein acts in the general stress response pathway of *B. subtilis*

One of the strongest and most obvious responses of *B. subtilis* cells to a wide range of stress and starvation conditions is the induction of the so-called general stress regulon, consisting of more than 150 genes/proteins [96] that are selectively expressed or regulated in *B. subtilis* in order to cope with changing environmental situations. Those stimuli can include high and low temperature, salt stress, ethanol- and acid stress, as well as cell wall stress imposed by the addition of antibiotics such as vancomycin and bacitracin (environmental signaling branch). The second signaling branch facilitates the response to starvation for glucose, phosphate, and oxygen (energy stress branch) [96]. The expression of the whole regulon is primarily regulated by controlling the activity of the master regulator, σ^B an alternative sigma factor that recognizes a particular promoter structure [97] and hence initiates gene expression of the regulon members. The aforementioned environmental branch of this pathway contains a family of five paralogous proteins that function either as negative regulators (RsbRA, RsbRB, RsbRC, and RsbRD) or as positive regulator (YtvA) of σ^B [91, 98]. Several proteins, including those five paralogous proteins, form together with RsbS (an inhibitor of the positive σ^B regulator RsbT) a large environmental signaling complex (also referred to as the Stressosome [99]) that controls the activation of σ^B via the above mentioned environmental signaling branch [91, 96]. All five paralogous proteins as well as RsbS contain a STAS domain, hence probably act in an antagonist / co-antagonist manner controlling σ^B activity [91, 98] via the control of the partner-switching modules RsbT-RsbS. For a comprehensive review of the complex σ^B activation network in *B. subtilis*, see Hecker (2007) [96]. The LOV domain-containing RsbS paralog YtvA, which was only recently included in the Stressosome signaling complex and hence in the environmental signaling branch, was shown to be directly involved in a blue-light dependent enhancement of the σ^B response [90-92]. Furthermore, the photoactive cysteine residue of YtvA (Cys62) was necessary to facilitate the YtvA specific response [91, 92]. When Cys62 (or in general the cysteine residue of the canonical sequence motif, NCRFLQG) is mutated to alanine or serine, the signaling state intermediate of the LOV photocycle can not be formed, thus abolishing the LOV photocycle in YtvA and in other LOV

domains [64, 100]. Hence, the conformational changes that may accompany the signaling state formation in YtvA are necessary to elicit its physiological function.

1.7.2 Blue-light activated LOV histidine kinases

Briggs, Bogomolni and co-workers very recently characterized a new bacterial protein family of histidine kinase associated LOV-photoreceptors (LOV-HKs) that are found in several proteobacterial species, namely in the plant pathogen *P. syringae* pathovar tomato, the marine phototroph *E. litoralis* and in two human/animal pathogens *B. melitensis* and *B. abortus*. For the latter member of the family, the *B. abortus* LOV-HK (BA-LOV-HK), it could be demonstrated that the corresponding gene is required for optimal replication and thus survival of *B. abortus* in murine macrophages, clearly suggesting that the BA-LOV-HK photoreceptor serves as a virulence factor in *B. abortus* [83].

Furthermore, the authors could demonstrate that the reactive cysteine residue (C69) in the LOV domain of BA-LOV-HK is strictly required for the mediation of the observed blue-light effect, thus confirming that the conformational rearrangement in the LOV domain of the protein, which is triggered by blue-light, initiates the final physiological response. It was speculated that, when *Brucella* is aborted as part of the placenta of the infected animal, the organism suddenly encounters a dramatically changed environment (outside of its host). In this environment the suddenly present light stimulus might prepare the bacteria for the infection of the next host [83].

1.8 Scope and outline of this thesis

At the onset of this study, the available information regarding bacterial LOV domain-containing proteins was, with respect to the photochemistry, limited to the *B. subtilis* YtvA protein and to the LOV domain of a *C. crescentus* LOV histidine kinase. Information regarding their biological function in any bacterial host was furthermore virtually non-existent. Therefore, this study was initiated to broaden the knowledge about bacterial LOV domain-containing proteins in general, thus including aspects such as phylogenetic inheritance, biochemical / biophysical characterization of novel bacterial LOV proteins, as well as their biological significance for the respective host organism.

The main part of the presented thesis consists of the attached publications and is additionally completed by so far unpublished data to strengthen the arguments in the published work. The content brought together here highlights the different aspects of the same general scientific theme: bacterial blue-light photoreceptors of the LOV-family. This main part is divided into three different sub-chapters that extend from Chapter 2 to 4 respectively.

Chapter 2 deals with evolutionary processes that contributed to the emergence of the four LOV domain subfamilies that we know today: i) bacterial LOV proteins, ii) eukaryotic (plant) phototropin-LOVs, iii) the ZTL/ADO/FKF1-LOV family and iv) the fungal white-collar-1 LOV family.

Chapter 3 focuses on the bacterial LOV domain-containing proteins especially corroborating eventually shared signal-transduction mechanisms between the extensively studied eukaryotic LOV signaling systems and their counterparts in the prokaryotic world. Furthermore, the chapter includes so far unpublished data on a possible biological function for one set of those bacterial LOV domain-containing signaling proteins, the *P. putida* PpSB1-LOV and PpSB2-LOV system.

Chapter 4 describes biotechnological applications of the LOV photoreceptor module and points toward possible future directions in this emerging field of research.

2 Distribution and evolution of the LOV domain signaling module

2.1 LOV is all around - A phylogenetic study on the origin of LOV-dependent blue-light signaling

Krauss U., Minh B.Q., Losi A., Gärtner W., Eggert T., von Haesler A. and Jaeger K. E.

Submitted for publication

SUPPLEMENTARY TABLE 1: DETAILED LOV SEQUENCE INFORMATION (SEQUENCE-IDs, ACCESION NUMBERS, DOMAIN CONTENT) CAN BE FOUND IN APPENDIX C

LOV is all around – A phylogenetic study on the origin of LOV-dependent blue-light signaling proteins

Ulrich Krauss[1], Bui Quang Minh[4], Aba Losi[2], Wolfgang Gärtner[3], Thorsten Eggert [5], Arndt von Haeseler[4] and Karl-Erich Jaeger[1]

Submitted for Publication

Fungi and plants are known to respond to environmental light stimuli via the action of different photoreceptor modules. One such class, responding to the blue-region of light, is constituted by the photoreceptors containing so-called light oxygen voltage (LOV) domains as light sensor modules. Three distinct families are identified in eukaryotes: i) the plant PHOTOTROPINS responsible for physiological effects such as plant phototropism, chloroplast relocation and stomatal opening in response to blue-light stimuli, ii) the plant circadian photoreceptors of the ZEITLUPE(ZTL)/ADAGIO(ADO)/Flavin-binding-Kelch-Repeat F-Box protein1 (FKF1) and iii) the fungal circadian photoreceptor WHITE-COLLAR-1 (WC-1). Interestingly, blue-light sensitive LOV signaling modules are widespread also throughout the prokaryotic world. In this study, we investigate the evolutionary scenarios for the appearance of LOV-based blue-light sensors in the eukaryotic and prokaryotic kingdoms of life. The phylogenetic analysis is based on 128 putative LOV domain sequences, representing all currently identified prokaryotic and a selection from the three major eukaryotic LOV domain photosensor families. Our results suggest a prokaryotic origin of the three eukaryotic LOV-photoreceptor families. Moreover, we highlight the importance of endosymbiosis, gene duplications as well as domain shuffling (fusion and fission) in the evolutionary process that apparently resulted in the wide distribution of the LOV light-signaling sensory systems within the three kingdoms of life.

Introduction

Phototropism in plants, defined as the movement or elongation of plant hypocotyls towards a unidirectional source of light, is a well-known phenomenon. It was first scientifically described by Charles Darwin in his book dating back to 1880: "The power of Movement in plants" (Darwin 1880). Nevertheless, it took more than a century until the photoreceptor mediating these responses was identified by Winslow Briggs and colleagues (Huala et al. 1997). The gene *nph1* (non-phototrophic hypocotyl) whose product was later named phototropin (phot) (Christie et al. 1999) was identified in forward genetic screens in *Arabidopsis thaliana* (Liscum and Briggs 1995) to act very early in the signaling cascade of the blue-light dependent phototrophic response (Huala et al. 1997).

After the identification of phototropins it soon became apparent that other blue-light driven phenomena in plants, such as chloroplast-relocation, stomatal opening as well as leave opening and expansion are mediated by the very same photoreceptor (Kagawa et al. 2004; Ohgishi et al. 2004).

Likewise, the phototropin of the green alga *Chlamydomonas reinhardtii* is involved in blue-light mediated gametogenesis (Huang and Beck 2003). In the eukaryotic phototropin systems, two so-called light, oxygen, voltage (LOV) photosensor domains, LOV1 and LOV2 were identified as the sensor switches (Christie et al. 1999) of a phosphorelay-signaling cascade which initiate the above mentioned physiological responses whereas a serine/threonine kinase located C-terminally to the LOV sensor domains is autophosphorylated in a blue-light dependent manner (Christie et al. 1998).

Another class of eukaryotic LOV-domain containing proteins is constituted in particular by the ZTL and the FKF1 proteins (ZTL/ADO/FKF1-family). Those were found to play a primary role in the photocontrol of flowering in plants (Nelson et al. 2000) and the light-dependent regulation of the circadian period in higher plants (Somers et al. 2000). The best characterized fungal LOV-domain containing protein is the white-collar 1 (WC-1) protein of *Neurospora crassa* (Ballario et al. 1996), that is involved in the blue-light dependent control of circadian responses in this fungus (Froehlich et al. 2002). Moreover, *Neurospora crassa* WC-1 was identified as the key regulator of all blue-light responses, namely induction of mycelial carotenogenesis, promotion of conidia and protoperithecia development, phototropism of perithecial beaks and inhibition or shifts in the circadian rhythm of conidiation (see (Ballario et al.

[1] *Institut für Molekulare Enzymtechnologie, Heinrich-Heine Universität Düsseldorf, Forschungszentrum Jülich, D-52426 Jülich, Germany*
[2] *Dept. of Physics, University of Parma,I-43100-Parma, Italy*
[3] *Max-Planck-Institut für Bioanorganische Chemie, Stiftstr. 34-36, D-45470 Mülheim, Germany*
[4] *Center for Integrative Bioinformatics Vienna, Max F. Perutz Laboratories, University of Vienna, Medical University Vienna, University of Veterinary Medicine Vienna Dr. Bohr Gasse 9, A-1030 Vienna, Austria*
[3] *EVOcatal GmbH, Merowingerplatz 1a, D-40225 Düsseldorf, Germany*

Supplementary Table 1: LOV domain containing proteins of the study
Supplementary Figure 1: Multiple sequence alignment of LOV-domains
Supplementary Figure 2: Alignment of 16S rRNA gene sequences

1996) and references therein). With the increasing number of completely sequenced microbial genomes it became evident that many bacteria are equipped with a variety of putative photosensory proteins such as red/ far-red sensing phytochromes (Davis, Vener, and Vierstra 1999), rhodopsins (Beja et al. 2000), and several blue-light sensing photoreceptors such as cryptochromes (Brudler et al. 2003), and sensors of the blue-light sensing using FAD (BLUF) (Gomelsky and Klug 2002) and light, oxygen, voltage (LOV) (Losi et al. 2002; Crosson, Rajagopal, and Moffat 2003) photoreceptor families. The latter LOV sensory modules that were previously identified as the key modulators of a myriad of plant and fungal blue-light responses, were surprisingly found equally widespread in photosynthetic and non-photosynthetic prokaryotic organisms (Losi 2004). Thus, the major light sensing paradigms are shared between such distantly related taxa as fungi, plants, bacteria and archaea.

Biochemical/photochemical data suggest that prokaryotic LOV-domain containing proteins undergo a photochemistry similar to the LOV domains of plant phototropins (phot) (Losi et al. 2002; Losi 2004; Krauss et al. 2005; Narikawa et al. 2006; Briggs 2007). This alone, as well as the sheer abundance of these sensor modules among prokayotes suggests blue-light driven physiological responses in their corresponding bacterial/archaeal hosts.

Although experimental evidence for archaeal LOV domains with the same photochemistry (as in bacteria and eukaryotes) is still missing, sequence analysis and the conservation of the mechanistically crucial amino acids strongly suggests a conservation of the LOV photosensing paradigm also in the archaeal lineages (Losi 2004; Losi 2006).

On the molecular level, the light sensitive function of the LOV-proteins is strictly based on the presence of a cysteine residue located in a distance of about 4 Å to the isoalloxazine ring of the flavin chromophore. Upon irradiation, a covalent bond is formed between this cysteine and the carbon-atom in position 4a of the flavin isoalloxazine ring, which thermally opens again within minutes to hours, dependent on the LOV protein (Losi 2006; Matsuoka et al. 2006). In this article, we will refer only to phot-like LOV domains bearing the reactive cysteine (highlighted bold-underlined) in the canonical LOV sequence motif GXN**C**RFLQG and therefore presumably exhibiting phot-LOV like photochemistry.

In the last two years it was demonstrated in several ground-breaking studies that blue-light is indeed an important environmental stimulus even in non-photosynthetic prokaryotes (Avila-Perez, Hellingwerf, and Kort 2006; Gaidenko et al. 2006; Swartz et al. 2007).

So far blue-light dependent physiological responses are described for a two component signaling system in the common soil bacterium *Bacillus subtilis.* In *B. subtilis* the YtvA protein, having been the first photochemically characterized LOV sensor module (Losi et al. 2002), shows a phot-like photochemistry and acts in a complex signaling network that regulates the general stress response pathway in *B. subtilis* in a blue-light dependent manner (Avila-Perez, Hellingwerf, and Kort 2006; Gaidenko et al. 2006; Suzuki et al. 2007). Very recently, Bogomolni and co-workers were able to demonstrate that for the intracellular animal pathogen *Brucella abortus,* efficient cell-replication in its host and hence virulence is strongly stimulated by blue-light (Swartz et al. 2007). *Brucella abortus* as well as two other *Brucella* species, among those also the notorious human pathogen *Brucella melitensis,* harbour LOV domain containing histidine kinases. The authors could unequivocally demonstrate that the LOV-domain containing kinase is blue-light sensitive, possessing a photochemistry similar to the plant phot-system (Swartz et al. 2007). Furthermore, like in plant phototropins, irradiation causes an autophosphorylation in the LOV associated kinase that triggers the above mentioned physiological response (Swartz et al. 2007). A similarly structured system, consisting of a LOV-domain, a histidine kinase and a fused response regulator has recently been identified for the plant pathogen *Pseudomonas syringae* pv tomato (Cao et al. 2007)

Finally, only several years after the identification of the plant phot-system with its blue-light-sensitive LOV sensory modules, experimental evidence accumulates for similar systems in bacteria, namely LOV histidine kinases, which can regulate blue-light dependent responses via phot-like autophosphorylation systems. This raises questions regarding the evolutionary history of the LOV photosensor module in the pro- and eukaryotic domains of life. Currently, no in-depth and up-to-date phylogenetic analyses are available that include all putative LOV-domain representatives found in both pro- and eukaryotic taxa and thus questions regarding their evolution are still unresolved. We therefore performed a phylogenetic analysis that comprised 129 putative LOV sequences from 11 eukaryotic, 65 bacterial and 3 archaeal taxa.

Material and Methods

Sequence retrieval and domain content analysis

Protein sequences with significant sequence similarity to LOV (Light Oxygen Voltage) domains were obtained using the PSI-BLAST (Altschul et al. 1997) utility of the NCBI homepage using several eukaryotic and bacterial LOV domain sequences of proteins with demonstrated function or photochemistry as query-sequences in independent BLAST searches. The resulting hits were manually filtered regarding the presence of the canonical LOV sequence motif GXN**C**RFLQG and, more importantly, of the photoreactive cysteine residue therein (highlighted in bold). The remaining sequences of the independent runs were subsequently used as input for a second PSI-BLAST iteration. Again sequences without the canonical sequence motif and furthermore duplicate hits from the independent BLAST runs were discarded.

The obtained list sequences covered representatives of divergent taxa and kingdoms, including eukaryotic phots (Briggs and Christie 2002), ZTLs (Somers et al. 2000), fungal WC-1 (Ballario et al. 1996) sequences as well as archaeal LOV domain containing proteins, and most importantly a comprehensive list of LOV homologs from a variety of bacterial taxa. Phototropin sequences were divided in LOV1 and LOV2 domains, and separately included in the alignment. The retrieved full-length protein sequences were subjected to a functional domain content analysis using the Simple Modular Architecture Research Tool (SMART)(Schultz et al. 1998; Ponting et al. 1999) at the European Molecular Biology Laboratory (EMBL) http://smart.embl-heidelberg.de/.
The 16s rRNA gene sequences that were used to reconstruct the phylogenetic species tree of the LOV photoreceptor protein containing taxa, were obtained either from the NCBI Genebank database (www.ncbi.nlm.nih.gov), from the GreenGenes 16s-rRNA gene database (http://greengenes.lbl.gov) (DeSantis et al. 2006), or from the Ribosomal Database Project II (http://rdp.cme.msu.edu/) (Cole et al. 2007). For the plant species in our analysis, the respective nuclear encoded small subunit ribosomal RNA gene sequences were used to reconstruct the three kingdoms species tree.

Sequence Alignment

Protein sequences of the LOV domain were aligned with M-Coffee (Wallace et al. 2006). M-Coffee is a meta alignment method which combines five different aligning strategies T-Coffee (Notredame, Higgins, and Heringa 2000), Clustalw (Thompson, Higgins, and Gibson 1994), Muscle (Edgar 2004), Mafft (Katoh et al. 2002), and Probcons (Do et al. 2005). Independent alignments were also constructed using these five methods to assess the quality of the final alignment. The DNA alignment of the LOV domain was subsequently obtained from the protein alignment based on the corresponding codon frames. The 16S rRNA sequences were aligned directly with M-Coffee and other methods using the same methodology. Resulting LOV and 16S alignments are available in the supplementary materials. Visual inspection and judgement based on alignment scores showed that M-Coffee produced the best result, although T-Coffee and Clustalw produce similar alignments. Hence, the alignment from M-Coffee was edited manually and used in subsequent phylogenetic analysis.

Phylogenetic Tree Reconstruction

Bayesian analysis
Six different runs of MrBayes (Huelsenbeck and Ronquist 2001) were conducted to account for a possible trapping in local optima. Due to a large number of sequences in both alignments we use one Markov Chain Monte Carlo (MCMC) chain per run, set the total number of generations to 50 millions and sampled every 1000 generation. At the end, each run produced a collection of 50,000 trees except the random starting tree. The convergence of all runs was then judged by the tool Tracer (http://tree.bio.ed.ac.uk/software/tracer/) (Rambaut and Drummond 2007).
The best run was finally used with a suitable burn-in to summarize the MrBayes tree.

For the LOV sequences analysis results showed that two of the six runs were trapped in sub-optima, suggesting that this is a difficult data set. However, the remaining four runs converged to the same plateau. The best run was identifed to converge after 10 million generations. The burn-in value was therefore set to 10,000. In contrast, the tree likelihood from all MCMC chains on the 16S gene converged very early to the same peak.

Bootstrap with maximum likelihood

A non-parametric bootstrap was used to evaluate the reliability of the tree topology. 1000 bootstrap samples from each original alignment were generated using the *seqboot* program from the PHYLIP package (Felsenstein 1989). From each bootstrap alignment we reconstructed a maximum likelihood (ML) tree using the IQPNNI program version 3.2 (Vinh and Von Haeseler 2004). For the IQPNNI heuristic, at least 200 iterations were executed and the stopping rule was turned on to automatically determine how many iterations are needed to reach a ML tree with 95% confidence. The number of iterations was limited to at most 2000 to avoid the program from running too long. Subsequently, an extended majority-rule consensus tree was constructed using PHYLIP's *consense* program.

The HKY85+G model of substitution (Hasegawa, Kishino, and Yano 1985; Yang 1994) was used for the LOV domain data set, whereas the TN93+G model (Tamura and Nei 1993; Yang 1994) was used for the 16S rRNA gene data both with four discrete gamma rate categories. Note that the Modeltest program (Posada and Crandall 1998) suggested TVM+I+G for the LOV data and TN93+I+G for the 16S data. However, we decided to use a simpler evolution model for the LOV domain and the 16S rRNA gene data, because of the short sequence length and the estimated proportion of invariable sites that is not sufficiently high for both data sets.

Maximum likelihood tree

In order to determine the best ML tree, 30 independent runs of IQPNNI were started, all with stopping rule enabled and with no limit on the number of iterations. The tree with highest likelihood among all runs was considered as the ML tree.

Evaluation of tree topologies

In order to determine whether MrBayes, bootstrap and maximum likelihood trees are significantly different or not, several topology tests were performed, including the one-sided KH test (Kishino and Hasegawa 1989), SH test (Shimodaira and Hasegawa 1999), expected likelihood weight (ELW) (Strimmer and Rambaut 2002), and the approximately unbiased test (AU; (Shimodaira 2002). The significance level of

all tests is set at 0.05. The tests were performed with Tree-puzzle (Schmidt et al. 2002) and CONSEL (Shimodaira and Hasegawa 2001)

Results and Discussion

LOV is all around

Previous studies had indicated that the LOV signaling module is widely distributed in all three kingdoms of life but apparently is absent in animals (Losi 2004). A BLAST analysis using PSI-BLAST (Altschul et al. 1997) (www.ncbi.nlm.nih.gov) as well as independent genome mining in various fungal genome projects (accessible at the Munich Information Center for Protein Sequences, http://mips.gsf.de) was performed to retrieve sequences with significant similarity to bacterial and eukaryotic LOV domains. The PSI-BLAST results were manually filtered regarding the presence of the canonical LOV sequence motif GXN**C**RFLQG and more importantly for the presence of the photoreactive cysteine residue therein (highlighted in bold).
These analyses revealed that the LOV signaling module occurred in about 12% of bacterial taxa (65 organisms out of 521 completely sequenced genomes) and correspondingly in approximately 7% of archaeal taxa (3 organisms out of 46 completely sequenced genomes) (Entrez-genome, September 2007). Among the Archaea, LOVs seem to be restricted to a few haloarchaeal lineages, whereas in the bacterial kingdom they are much more widely distributed. They are widely dispersed throughout the Proteobacteria, are represented among the Firmicutes (esp. *Bacillus* and *Listeria* genera) and the Cyanobacteria and were so far found in only a few Chloroflexi and Actinobacteria. It should be noted that the rapidly growing number of sequenced bacterial genomes as well as the huge amount of metagenome sequences deliver novel LOV sequences almost monthly. Therefore, a sequence analysis can *de facto* never be complete but merely represents a snapshot of the sequence and species data available at the given moment. Additional to the prokaryotic LOVs for which the sequence list should be complete for the time being, representative sequences for the currently known three major eukaryotic LOV-photoreceptor families, namely phototropins-LOVs, ZTL/ADO/FKF1-LOVs and the fungal WC-1 family, were included in the analysis to gain insight into evolutionary relationship between the prokaryotic and the eukaryotic LOV-systems.

Tree reconstruction and evaluation

The phylogenetic analysis presented in this study was performed exclusively on the conserved LOV core domain (the blue-light sensitive sensor domain) of the putative and known photoreceptor proteins. This choice was dictated by an overwhelming diversity of the functional domains found additionally to the LOV sensor domain within the photoreceptors in general and especially among the prokaryotic LOV-containing proteins. Therefore, no meaningful alignment can be generated using full-length photoreceptor sequence information. Thus, it is impossible and not meaningful to construct a tree for the three kingdoms of life based on full-length LOV photoreceptor sequences.

An initial tree reconstruction based on an amino acid (aa) sequence alignment of the LOV domains, produced a more or less un-resolved tree (star-like topology). This might be attributed to convergent evolution on the aa level, due to functional restriction. As this would obviously be less pronounced on the nucleotide level, due to the degenerate nature of the genetic code, we used the corresponding nucleotide sequence alignment of the LOV core domain throughout this study. Interestingly, all 30 different runs of IQPNNI, carried out to identify the best ML tree, of the LOV sequences, reconstructed 30 different tree topologies, whereas on the 16S rRNA gene all runs inferred the same tree. This again indicates that the LOV alignment is a difficult data set and more runs are needed to gain more confidence about the ML tree.

Several tree topology test were performed on the MrBayes, bootstrap and the ML tree reconstructed for the LOV domain data set (see materials and methods section for details). All tests except ELW indicated that no significant topological difference between trees is observed, whereas the ELW test rejected the bootstrap consensus tree only with a marginal p-value of 0.03.
Therefore, the phylogenetic tree generated by MrBayes is used for the illustration throughout this study. For the 16S rRNA gene all three trees are very similar and thus no tests were performed.

We have to note here, that even with the use of the nucleotide alignment the obtained tree resolution is low and the phylogentic signal is weak in particular at certain deep branches and in the bootstrapped tree. This might in general be attributed to the short sequence length, and thus to the low amount of data that is available to reconstruct the evolution.
Nevertheless, the branches that we are going to discuss in the following in more detail are among the ones with
the highest support.

Topology of the three kingdoms phylogeny

The tree shown in Fig. 1 clearly supports a separation of bacterial and archaeal LOV sequences with a strong evolutionary signal. This prokaryotic dichotomy is well in accordance with the 16S rRNA gene tree (Fig. 2) that comprises the same taxa as the LOV domain tree. Therefore, the tree shown in Fig.1 was rooted using the archaeal LOV sequences.

We observe a separation into two subtrees in the non-archaeal part of the tree, the top grouping consists of bacterial sequences including plant sequences

(ZTL/ADO/FKF1-LOV), whereas the lower part is a mixture of eukaryotic sequences from fungi (WC-1-LOVs) and plants (phototropin-LOVs), and of protobacteria. This dichotomy does not occur in the 16S-rRNA tree.
The evolutionary events which may have resulted in the distribution and appearance of LOV among the three kingdoms of life are illustrated in Fig. 3. In the bacterial subtree of the LOV tree, the general topology is by large congruent with the topology of the ribosomal tree. For example the Firmicutes and Cyanobacteria are clearly separated from the majority of the Proteobacteria (α,β,γ). Interestingly, the different eukaryotic LOV photoreceptor families (with the exception of the fungal sequences) are monophyletic, whereas the respective clades show a clear affinity to certain bacterial lineages and are not separated from bacteria and archaea as observed in the 16S rRNA tree.

The bacterial origin of the eukaryotic LOV signaling systems

Regarding the evolutionary history of the LOV domain photoreceptor module and hence, of LOV-dependent blue-light perception in the plant and fungal kingdoms, the phylogenetic tree presented in Fig. 1 points in all cases toward a direct sister-group relationship between bacterial LOVs and certain eukaryotic LOV lineages.
Firstly, the eukaryotic ZTL/ADO/FKF1-LOV family clusters within the cyanobacterial LOV sequences (Bayesian support 0.79), but it does not form a monophyletic group together with the other eukaryotes (the fungi and phot-LOVs), as could be expected under the classic view of evolution. Thus, the LOV domain of the ZTL/ADO/FKF1-LOV family in higher plants probably was already present in the photosynthetic cyanobacterial endosymbiont which gave rise to the eukaryotic chloroplast.
Secondly, the remaining two eukaryotic LOV photoreceptor families, namely the plant phototropin-LOVs and the LOV domains of the fungal circadian photoreceptor WC-1, are related with an α-proteobacterial clade, which comprises mainly marine phototrophic α-proteobacteria, including among other species *Rhodobacter spaeroides* (Alpha3), *Roseobacter dentrificans* (Alpha2_1), *Dinoroseobacter shibae* (Alpha6) and *Erythrobacter litoralis* (Alpha8_1-4). This suggests also in case of the latter two eukaryotic LOV photoreceptor families the involvement of endosymbiosis (here with an ancient proteobacterium) in the early evolutionary processes that led to the appearance of LOV in the eukaryotes. Because the Bootstrap and Bayesian support values for this grouping are weaker (0.63), we can however not exclude the possibility that this part of the tree follows the classical view of three kingdoms. Notably, some of the ZTL sequences (that cluster with the cyanobacterial clade) and the phototropin LOV sequences (that show a clear affinity towards the proteobacterial groups) are even found within the nuclear genome of the same species, namely *Arabidopsis thaliana*.

The apparent scenario for the appearance of LOV in the eukaryotes therefore is well in accordance with the general endosymbiotic theory (Mereschkowsky 1905; Wallin 1927). Today, most biologists agree that an α-proteobacterium (Esser, Martin, and Dagan 2006), probably belonging to an ancient photosynthetic (non)-sulphur bacterial lineage (Cavalier-Smith 2006), was the ancestor of eukaryotic mitochondria. The chloroplasts, on the other hand, are thought to have originated from the endosymbiosis of an ancestral cyanobacterium (McFadden 2001). Therefore, the evolutionary scenario presented for the LOV domain containing photoreceptors and thus the general topology of the LOV gene tree seem feasible in the light of the current view of endosymbiotic theory whereas the general topology of the species tree generated based on 16s rRNA gene sequences follows the classic view of the three kindoms.

An observation challenging the endosymbiotic hypothesis for the appearance of LOV among the eukaryotes relates to the exclusive localization of the eukaryotic LOV photosensors in the nuclear but not the chloroplast or mitochondrial genomes of the respective plants. However, this discrepancy might be accounted for by invoking gene transfer from organelles to the nucleus after the endosymbiotic uptake event. This process which is known as endosymbiotic gene transfer is widely accepted today and is comprehensively reviewed by Timmis, Martin and co-workers (Timmis et al. 2004).
It is interesting to note that both ZTL/ADO/FKF1, the phototropins and WC-1 LOVs probably originated directly from bacterial species that were able to perform photosynthesis, either oxygenic as in case of the cyanobacteria or anoxygenic as in case of the α-proteobacteria of our analysis. In light of this observation it seems feasible to suggest that photosynthesis, or more generally, phototrophy, defining a metabolic mode in which organisms convert light energy into chemical energy for growth (Bryant and Frigaard 2006) might have been a prerequisite for the "invention" of a blue-light sensor in prokaryotes. This definition also includes the Archaea as halophilic but also phototrophic organisms capable to sustain metabolic processes by means of light-driven proton pumps (bacteriorhodopsins) (Haupts, Tittor, and Oesterhelt 1999; Lanyi 2004).

LOV domain distribution among the three domains of life may have involved duplication events

In Fig. 3 the evolutionary scenarios that might have contributed to the wide distribution of LOV among the three kingdoms are summarized. The affinity of certain (α, β, γ) - proteobacterial genera to the two major branches of the non-archaeal LOV gene tree (Fig. 1) is inconsistent both with the tree describing general microbial phylogeny and with the 16S rRNA gene tree (Fig. 2). This fact might be explained by an ancient duplication event that took place before the separation of Firmicutes and Cyanobacteria from the Proteobacteria occurred.

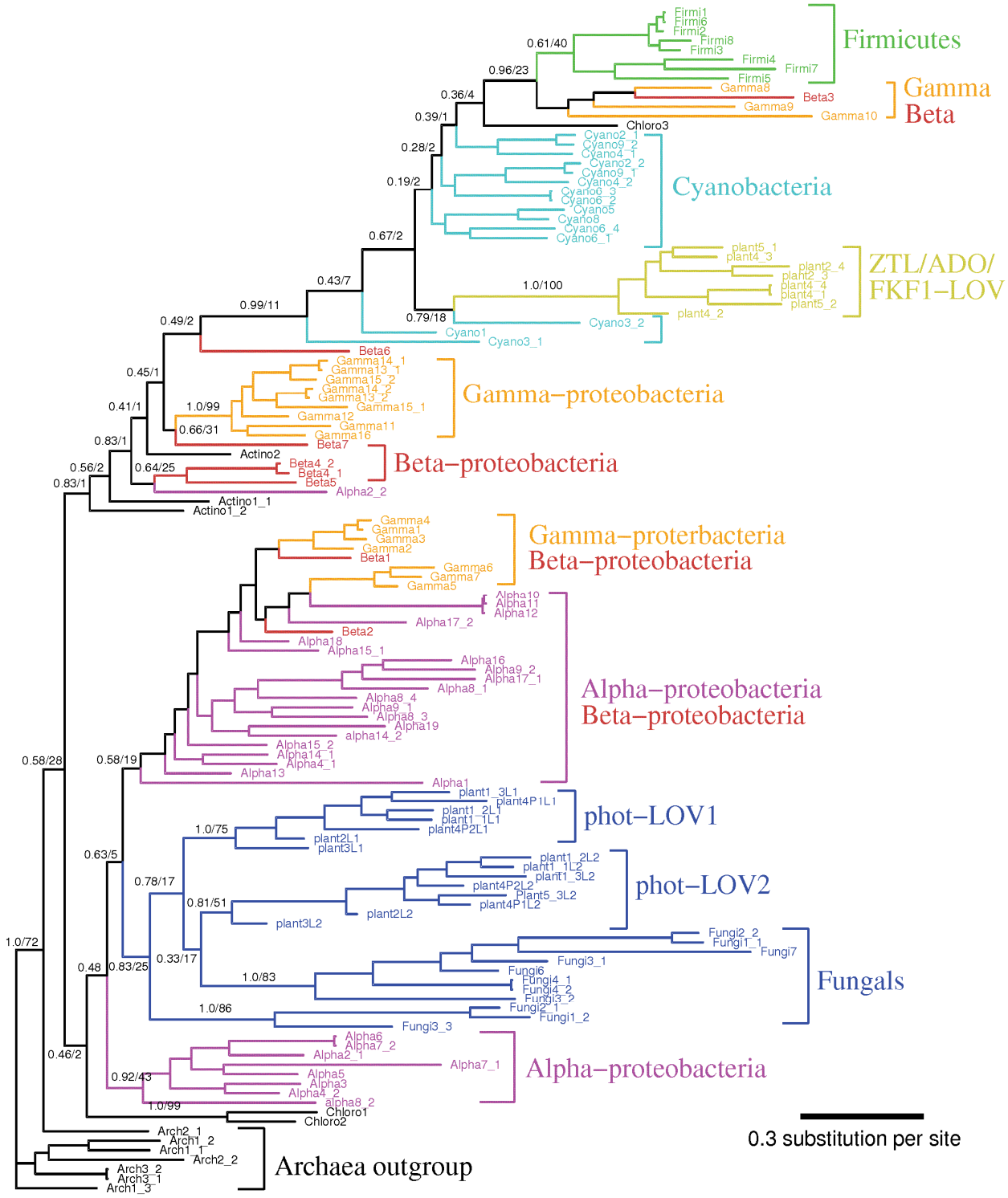


Figure 1: Phylogenetic tree, reconstructed for the LOV gene sequences from divergent taxa. Bayesian analysis and tree reconstruction was performed using MrBayes (Huelsenbeck and Ronquist 2001).Six independent runs were conducted to account for a trapping in local optima during the Markov Chain Monte Carlo (MCMC) run. After evaluation of the respective runs for convergence of the MCMC chain the best run was finally used with a suitable burn-in to summarize the MrBayes tree. Bayesian posterior probabilities and bootstrap (in this order) support values, based on 1000 replicates, are shown at the relevant branches of the MrBayes tree. The respective LOV sequences of the three different eukaryotic photoreceptor families as well as the LOV sequences of prokaryotic origin are color-coded: The Archaea as outgroup are shown in black,.α-proteobacteria in pink, β-proteobacteria in red and γ-proteobacteria in orange. Cyanobacterial genera are shown in cyan and the Firmicutes sequences are highlighted in dark green. Fungal WC-1 and plant-phot LOV (LOV1 and LOV2) are depicted in blue and the ZTL/ADO/FKF1-LOV family of the higher plants is shown in light green. Detailed sequence information, including protein accession numbers, are summarized in Table 1 in the Supplementary Material.

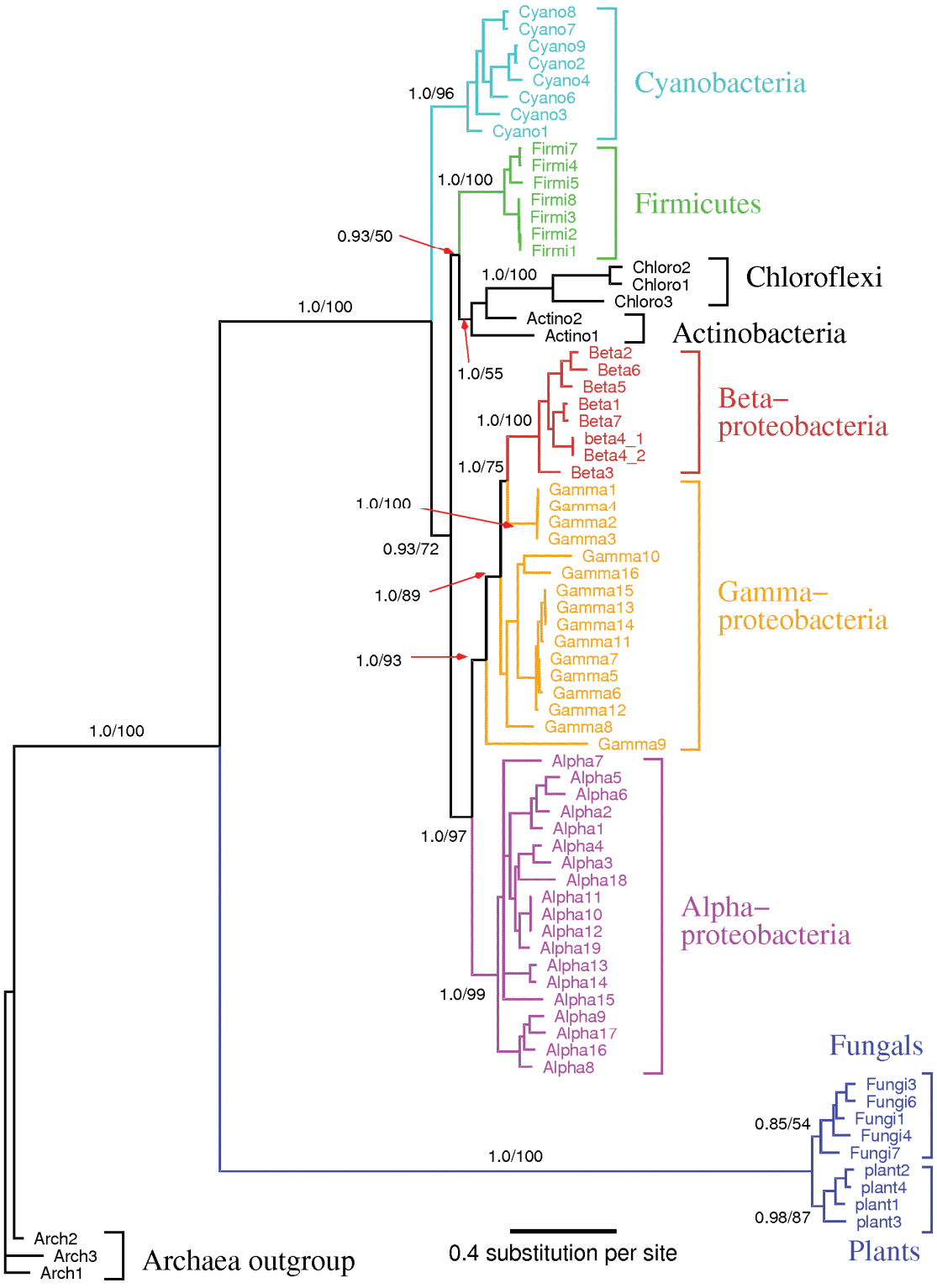


Figure 2: 16s rRNA gene tree reconstructed to infer the species evolution of the LOV-domain containing taxa. Bayesian, bootstrap and ML trees were constructed as described for the LOV gene tree. All three trees that were generated using the different phylogeny-inference methods showed the same overall topology. Here the MrBayes generated tree is depicted. Bayesian posterior probabilities and bootstrap support values are added at relevant branches. The sequences are color-coded in similarity to the LOV gene tree shown in Figure 1.

This assumption predicts subsequent gene loss in the Firmicutes and Cyanobacterial lineages to explain the absence of two phylogenetically distinct LOV domain subgroups in those genera. Another more recent duplication event in the eukaryotic LOV domain containing lineages (phot and WC-1) may have led to the presence of two LOV domains (LOV1 and LOV2) in plant phototropins.

A somewhat intriguing feature of the eukaryotic LOV phylogeny is the dichotomy of the fungal LOV sensor modules. Interestingly, they do not form a separate monophyletic fungal clade as would be expected under the assumption of the fungi as a late diverging eukaryotic group, but contrarily cluster in two phylogenetic groups (although the separation is only weakly supported) within the eukaryotic phot branch. However this feature might be explained by placing the duplication event into phototropin LOV1 and LOV2 well before the separation of the plant and fungal lineages.

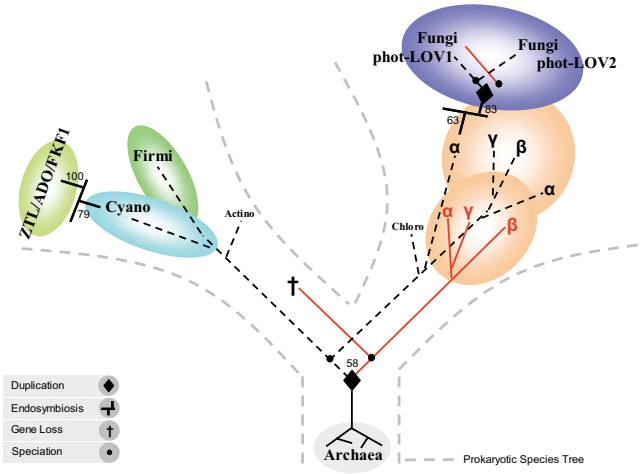


Figure 3: Illustration of our evolutionary hypothesis that may have led to the distribution of the LOV blue-light sensors in the three kingdoms of life. The figure illustrates the evolutionary scenarios that may have led to the wide dispersion of the LOV domain sensor module in the three kingdoms of life. The LOV signaling module might have its origin in phototrophic prokaryotes. An ancient duplication event in the bacterial branch of the LOV tree might have caused the proteobacterial dichotomy that is observed in the LOV gene tree, but is incongruent with the ribosomal (species) tree generated for the LOV domain containing taxa. A subsequent gene loss event is predicted to explain the absence of two distinct cyanobacterial / Firmicutes LOV sub-families. The prokaryotic LOV signaling module might have spread to the eukaryotes by endosymbiosis of ancient cyanobacteria (in case of the ZTL/ADO/FKF1 family (euZTL)) or proteobacteria (in case of phototropins (phot-LOV1 and phot-LOV2), respectively. Whereas two independent endosymbiotic events might explain the dichotomy of the plant LOV photosensors that is observed in our phylogenetic LOV gene tree, in contrast to the classical view of the three kingdom evolution represented in the ribosomal tree. Another more recent duplication event has to be invoked to account for the presence of two LOV domains (LOV1 and LOV2) in plant phototropins.

This hypothesis might even find support in the observation that the earliest diverging branch in both the distinct phot-LOV1 and LOV2 branches is the LOV1 (plant3_L1) and respectively LOV2 (plant3_L2) sequence of the phototropin from the green alga *Chlamydomonas reinhardtii*. Therefore, in agreement with the conclusion drawn by Lariguet and Dunand (Lariguet 2005), the *C. reinhardtii* phototropin might be close to the ancestral sequence from which plant phots have evolved. This furthermore implies that the duplication into phot-LOV1 and LOV2 clades must have occurred before the separation of the algae from the higher plants and hence *a priori* well before the separation of the fungal and plant eukaryotic lineages.

Bacterial LOV histidine-kinases as the ancestors of the plant phototropins

The phylogenetic tree reconstruction for the LOV photoreceptors was performed solely based on the conserved LOV core domain (i.e. the light sensor module), ignoring for the purpose of tree generation any additionally fused functional domains. Notably, the tree reconstructed from the LOV-domain multiple sequence alignment nevertheless results in a reasonable grouping of sequences with respect to their function, but also to the additionally fused functional domains of the respective full-length photoreceptor protein. An exemplary overview of the tremendous variety of full-length LOV photoreceptor multi-domain architectures is depicted in Fig. 4.
This multitude of architectures in which the LOV-domain is the only constant attribute clearly indicates that domain fission and fusion (Kummerfeld and Teichmann 2005; Cock and Whitworth 2007) must have played a primary role in the evolution of LOV domain containing photoreceptor proteins. In some cases, our phylogenetic analysis that was based only on the sensor LOV domain in these photoreceptors, might provide a basis for the deduction of domain acquisition (fusion) and loss (fission) events.

A striking example which we want to discuss in some more detail, is the appearance and distribution of LOV-kinase systems among the pro- and eukaryotes as illustrated in Fig. 5. In particular, the bacterial LOV-histidine kinases (LOV-HisKin) and hybrid LOV-histidine kinase (additionally containing a fused response regulator (RR)), sparked substantial interest lately, as they represent one of the few bacterial examples for which a blue-light dependent biological function (Cao et al. 2007; Swartz et al. 2007) has been established.

The evolutionary hypothesis we presented above predicts that the LOV domains of the plant phototropin photoreceptors probably originated from an endosymbiotic event of an ancient α-proteobacterium. Plant phototropins possess, apart from the two LOV domains, a fused serine/threonine kinase which is autophosphorylated in response to blue-light (Christie et al. 1998). Strikingly, the LOV proteins of the α-proteobacterial lineages, which we propose being direct ancestors of the plant phot-LOV systems predominately possess fused histidine kinases. For some of these systems a similar blue-light

dependent functionality, namely light-driven autophosphorylation was demonstrated (Cao et al. 2007; Swartz et al. 2007).

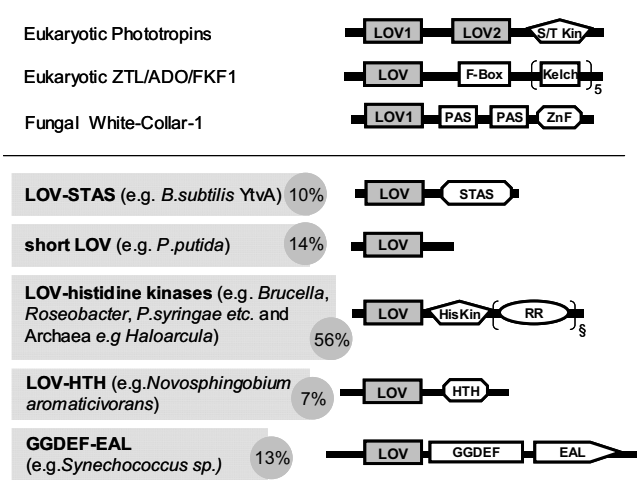


Figure 4: Exemplary multi-domain architectures of LOV photoreceptors. The domain content of full-length LOV-photoreceptor proteins was analyzed using the Simple Modular Architecture Research Tool (SMART)(Schultz et al. 1998; Ponting et al. 1999) at the European Molecular Biology Laboratory (EMBL) http://smart.embl-heidelberg.de/. Whereas the multi-domain LOV-architectures in the respective eukaryotic photosensor systems are completely conserved, the variety of full-protein organizations found in the prokaryotic domains of life is tremendous. Therefore, only a few paradigmatic LOV- photoreceptor architectures from the prokaryotes are depicted. The percentage values next to the respective group correspond to the proportion of this particular LOV-photoreceptor organization of the total prokaryotic sequences. Abbreviations: LOV: light, oxygen, voltage domain; S/T Kin: serine/threonine kinase; Kelch: Kelch repeats; PAS: Per, Arndt, Sim domain; ZnF: zinc-finger motif; STAS: sulphate-transporter antisigma-factor antagonist domain; HisKin: histidine kinase; RR: response regulator; HTH: helix-turn-helix DNA binding domain; GGDEF: diguanylate cyclase; EAL: phosphodiesterase; GAF: domain present in phytochromes and cGMP-specific phosphodiesterases;

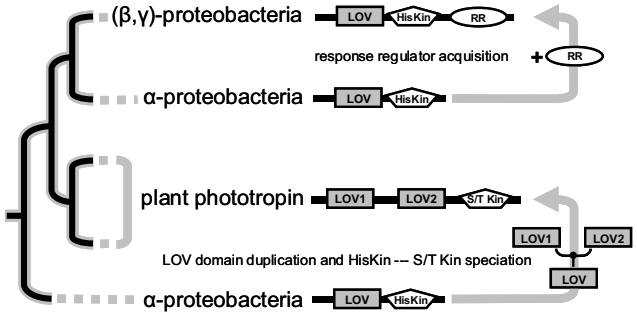


Figure 5: Distribution and appearance of LOV-kinase systems in the tree kingdoms of life. The figure illustrates how a LOV domain duplication and output domain speciation might have led to the emergence of today's phototropin system in the eukaryotes. A primordial LOV-histidine kinase from an ancient α-proteobacterium is suggested to be the ancestor of the eukaryotic phot system, whereas the uptake probably occurred through mitochondrial endosymbiosis. The same primordial bacterial systems (LOV-histidine kinases) might have further speciated, followed by a domain acquisition (fusion) event, hence resulting in today's hybrid LOV-histidine kinases. Interestingly this domain acquisition event coincides with a change in the microbial life-style, namely a shift from marine habitats (α-proteobacteria) towards a plant associated life style (symbiotic or pathogenic)(β,γ – proteobacteria).

We therefore propose that α-proteobacterial LOV-histidine kinases might represent the ancestors of the eukaryotic phototropins LOV-serine/threonine kinases. A similar scenario was previously suggested for the eukaryotic red/far-red light sensing phytochrome serine/threonine kinases that may possess a prokaryotic histidine kinase ancestry (Yeh and Lagarias 1998). Since also the few haloarchaeal LOV systems contain histidine-kinases fused to LOV sensor module, it is tempting to speculate that the primordial LOV photoreceptor systems were blue-light activated histidine-kinases.

The hybrid LOV-histidine kinases (LOV-HisKin-RR) are found nearly exclusively in bacterial plant pathogen species, e.g. *Xanthomonas spp.* (Gamma1-4) and *Pseudomonas syringae spp*. (Gamma5-6). Those systems probably derived from the ancestral α-proteobacterial LOV- histidine kinases by speciation (as judged from the phylogenetic tree in Fig. 1). Furthermore, a domain fusion event may account for the appearance of the fused response regulator in the hybrid LOV-histidine kinase architectures in those γ-proteobacterial plant pathogens (see Fig. 5). It is interesting to note that this domain acquisition event seems to coincide with a change in the lifestyle of the respective microbial species that contain those hybrid LOV-histidine kinases. All organisms containing the hybrid LOV-histidine kinases show a plant-associated or -symbiontic (e.g. *Bradyrhizobium spp* (Alpha18). and *Burkholderia phymatum* (Beta1) or a plant pathogenic (e.g. the above mentioned *Xanthomonas* and *P. syringae* species) life-style, whereas their probable ancestors the LOV-histidine kinases (lacking the RR) are found in predominately marine phototrophic α-proteobacteria (e.g. *Roseobacter* (Alpha6), *Fulvimarina* (Alpha3), and *Erythrobacter* (Alpha8) species etc.).

Conclusions

Based on the three kingdom phylogeny presented in this study we propose that the LOV signaling module represents an evolutionary ancient light sensor that retained its photochemistry from archaeal to bacterial and moreover to eukaryotic genera. The known eukaryotic LOV photosensor families may have originated by endosymbiosis from cyanobacterial or α-proteobacterial lineages. Consequently, the respective endosymbiotic uptake event should mark the time of the appearance of LOV among the eukaryotes. Moreover, our analysis suggests that the plant circadian photoreceptors of the ZTL/ADO/FKF1-LOV family and the respective plant phototropins have originated from two distinct endosymbiotic events separated by a few hundred million years in time. Hence, the distinct LOV photoreceptor families in plants (ZTL/ADO/FKF1-LOVs and phototropins) probably underwent independent divergent evolution toward distinct functions, namely light dependent regulation of i) the circadian rhythmicity and ii) phototropic responses. Contrarily, the circadian LOV photoreceptor systems found in plants (ZTL/ADO/FKF1-LOVs) and fungi (WC-1), which we

suggest to have originated from two different endosymbiotic events, convergently evolved to use the same light sensitive domain for the control of similar cellular processes, namely the light-entrainment of the circadian clock.

Apparently, the LOV photoreceptors retained the primary photochemical events that enable photon-capture and thus (blue)-light perception but are utilized for different responses in the three kingdoms of life. This is also reflected in the tremendous variety of additional functional domains that are found associated to the LOV sensor domain.

As an important result of our analysis, ancestral LOV-histidine kinases found today in marine phototrophic α-proteobacteria and in a few halophilic Archaea appear as among the evolutionary oldest and thus primordial LOV photoreceptor systems. Hence, phototrophy may have played a primary role in the evolutionary processes occurring in ancient marine habitats and has resulted in the "invention" of blue-light LOV photoreceptor systems which we find today dispersed in the three domains of life.

Acknowledgements

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) (Forschergruppe "Blue-Light Photoreceptors", FOR526). Financial support from the Vienna Science and Technology Fond (WWTF) is also greatly appreciated.

References

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**:3389-3402.

Avila-Perez, M., K. J. Hellingwerf, and R. Kort. 2006. Blue light activates the sigmaB-dependent stress response of Bacillus subtilis via YtvA. J Bacteriol **188**:6411-6414.

Ballario, P., P. Vittorioso, A. Magrelli, C. Talora, A. Cabibbo, and G. Macino. 1996. White collar-1, a central regulator of blue light responses in Neurospora, is a zinc finger protein. Embo J **15**:1650-1657.

Beja, O., L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich, and E. F. DeLong. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science **289**:1902-1906.

Briggs, W. R. 2007. The LOV domain: a chromophore module servicing multiple photoreceptors. J Biomed Sci **14**:499-504.

Briggs, W. R., and J. M. Christie. 2002. Phototropins 1 and 2: versatile plant blue-light receptors. Trends Plant Sci **7**:204-210.

Brudler, R., K. Hitomi, H. Daiyasu, H. Toh, K. Kucho, M. Ishiura, M. Kanehisa, V. A. Roberts, T. Todo, J. A. Tainer, and E. D. Getzoff. 2003. Identification of a new cryptochrome class. Structure, function, and evolution. Mol Cell **11**:59-67.

Bryant, D. A., and N. U. Frigaard. 2006. Prokaryotic photosynthesis and phototrophy illuminated. Trends Microbiol **14**:488-496.

Cao, Z., V. Buttani, A. Losi, and W. Gaertner. 2007. A blue light inducible two component signal transduction system in the plant pathogen Pseudomonas syringae pv. tomato. Biophys J.

Cavalier-Smith, T. 2006. Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. Proc Biol Sci **273**:1943-1952.

Christie, J. M., P. Reymond, G. K. Powell, P. Bernasconi, A. A. Raibekas, E. Liscum, and W. R. Briggs. 1998. Arabidopsis NPH1: a flavoprotein with the properties of a photoreceptor for phototropism. Science **282**:1698-1701.

Christie, J. M., M. Salomon, K. Nozue, M. Wada, and W. R. Briggs. 1999. LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. Proc Natl Acad Sci U S A **96**:8779-8783.

Cock, P. J., and D. E. Whitworth. 2007. Evolution of prokaryotic two-component system signaling pathways: gene fusions and fissions. Mol Biol Evol **24**:2355-2357.

Cole, J. R., B. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity, and J. M. Tiedje. 2007. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res **35**:D169-172.

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

Darwin, C. R. A. b. F. D. 1880. *The power of movement in plants*. John Murray, London.

Davis, S. J., A. V. Vener, and R. D. Vierstra. 1999. Bacteriophytochromes: phytochrome-like photoreceptors from nonphotosynthetic eubacteria. Science **286**:2517-2520.

DeSantis, T. Z., P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, and G. L. Andersen. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol **72**:5069-5072.

Do, C. B., M. S. Mahabhashyam, M. Brudno, and S. Batzoglou. 2005. ProbCons: Probabilistic consistency-based multiple sequence alignment. Genome Res **15**:330-340.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**:1792-1797.

Esser, C., W. Martin, and T. Dagan. 2006. The origin of mitochondria in light of a fluid prokaryotic chromosome model. Biol Lett.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics **5**:164-166.

Froehlich, A. C., Y. Liu, J. J. Loros, and J. C. Dunlap. 2002. White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. Science **297**:815-819.

Gaidenko, T. A., T. J. Kim, A. L. Weigel, M. S. Brody, and C. W. Price. 2006. The blue-light receptor YtvA acts in the environmental stress signaling pathway of Bacillus subtilis. J Bacteriol **188**:6387-6395.

Gomelsky, M., and G. Klug. 2002. BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. Trends Biochem Sci **27**:497-500.

Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol **22**:160-174.

Haupts, U., J. Tittor, and D. Oesterhelt. 1999. Closing in on bacteriorhodopsin: progress in understanding the molecule. Annu Rev Biophys Biomol Struct **28**:367-399.

Huala, E., P. W. Oeller, E. Liscum, I. S. Han, E. Larsen, and W. R. Briggs. 1997. Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain. Science **278**:2120-2123.

Huang, K., and C. F. Beck. 2003. Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga Chlamydomonas reinhardtii. Proc Natl Acad Sci U S A **100**:6269-6274.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754-755.

Kagawa, T., M. Kasahara, T. Abe, S. Yoshida, and M. Wada. 2004. Function analysis of phototropin2 using fern mutants deficient in blue light-induced chloroplast avoidance movement. Plant Cell Physiol **45**:416-426.

Katoh, K., K. Misawa, K. Kuma, and T. Miyata. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res **30**:3059-3066.

Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J Mol Evol **29**:170-179.

Krauss, U., A. Losi, W. Gartner, K. E. Jaeger, and T. Eggert. 2005. Initial characterization of a blue-light sensing, phototropin-related protein from Pseudomonas putida: a paradigm for an extended LOV construct. Phys Chem Chem Phys **7**:2804-2811.

Kummerfeld, S. K., and S. A. Teichmann. 2005. Relative rates of gene fusion and fission in multi-domain proteins. Trends Genet **21**:25-30.

Lanyi, J. K. 2004. Bacteriorhodopsin. Annu Rev Physiol **66**:665-688.

Lariguet, P., Dunand, C. 2005. Plant Photoreceptors: Phylogenetic Overview. J Mol Evol **61**:559-569.

Liscum, E., and W. R. Briggs. 1995. Mutations in the NPH1 locus of Arabidopsis disrupt the perception of phototropic stimuli. Plant Cell **7**:473-485.

Losi, A. 2004. The bacterial counterparts of plants phototropins,. Photochem Photobiol Sci **3**:566-574.

Losi, A. 2006. Flavin-Based Photoreceptors in Bacteria *in* E. Silva, Edwards, A.M.,, ed. Flavins Photochemistry and Photobiology. The Royal Soceity of Chemistry, Cambridge.

Losi, A., E. Polverini, B. Quest, and W. Gartner. 2002. First evidence for phototropin-related blue-light receptors in prokaryotes. Biophys J **82**:2627-2634.

Matsuoka, D., T. Iwata, K. Zikihara, H. Kandori, and S. Tokutomi. 2006. Primary Processes During the Light-signal Transduction of Phototropin. Photochem Photobiol.

McFadden, G. I. 2001. Chloroplast origin and integration. Plant Physiol **125**:50-53.

Mereschkowsky, C. 1905. Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol. Centralbl, **25**:593-604, 689-691.

Narikawa, R., K. Zikihara, K. Okajima, Y. Ochiai, M. Katayama, Y. Shichida, S. Tokutomi, and M. Ikeuchi. 2006. Three putative photosensory light, oxygen or voltage (LOV) domains with distinct biochemical properties from the filamentous cyanobacterium Anabaena sp. PCC 7120. Photochem Photobiol **82**:1627-1633.

Nelson, D. C., J. Lasswell, L. E. Rogg, M. A. Cohen, and B. Bartel. 2000. FKF1, a clock-controlled gene that regulates the transition to flowering in Arabidopsis. Cell **101**:331-340.

Notredame, C., D. G. Higgins, and J. Heringa. 2000. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol **302**:205-217.

Ohgishi, M., K. Saji, K. Okada, and T. Sakai. 2004. Functional analysis of each blue light receptor, cry1, cry2, phot1, and phot2, by using combinatorial multiple mutants in Arabidopsis. Proc Natl Acad Sci U S A **101**:2223-2228.

Ponting, C. P., J. Schultz, F. Milpetz, and P. Bork. 1999. SMART: identification and annotation of domains from signalling and extracellular protein sequences. Nucleic Acids Res **27**:229-232.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14**:817-818.

Rambaut, A., and A. J. Drummond. 2007. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18**:502-504.

Schultz, J., F. Milpetz, P. Bork, and C. P. Ponting. 1998. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A **95**:5857-5864.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol **51**:492-508.

Shimodaira, H., and M. Hasegawa. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**:1246-1247.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. . Mol. Biol. Evol. **16,** :1114-1116

Somers, D. E., T. F. Schultz, M. Milnamow, and S. A. Kay. 2000. ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. Cell **101**:319-329.

Strimmer, K., and A. Rambaut. 2002. Inferring confidence sets of possibly misspecified gene trees. Proc Biol Sci **269**:137-142.

Suzuki, N., N. Takaya, T. Hoshino, and A. Nakamura. 2007. Enhancement of a sigma(B)-dependent stress response in Bacillus subtilis by light via YtvA photoreceptor. J Gen Appl Microbiol **53**:81-88.

Swartz, T. E., T. S. Tseng, M. A. Frederickson, G. Paris, D. J. Comerci, G. Rajashekara, J. G. Kim, M. B. Mudgett, G. A. Splitter, R. A. Ugalde, F. A. Goldbaum, W. R. Briggs, and R. A. Bogomolni. 2007. Blue-light-activated histidine kinases: two-component sensors in bacteria. Science **317**:1090-1093.

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of

mitochondrial DNA in humans and chimpanzees. Mol Biol Evol **10**:512-526.

Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673-4680.

Timmis, J. N., M. A. Ayliffe, C. Y. Huang, and W. Martin. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet **5**:123-135.

Vinh, S., and A. Von Haeseler. 2004. IQPNNI: moving fast through tree space and stopping in time. Mol Biol Evol **21**:1565-1571.

Wallace, I. M., O. O'Sullivan, D. G. Higgins, and C. Notredame. 2006. M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res **34**:1692-1699.

Wallin, I. E. 1927. Symbionticism and the Origin of Species. Bailliere, Tindall & Cox, London.

Yang, Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol **39**:306-314.

Yeh, K. C., and J. C. Lagarias. 1998. Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry. Proc Natl Acad Sci U S A **95**:13976-13981.

3. Bacterial LOV proteins: from signal-transduction to physiology

3.1 Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct.

Krauss U., Losi A., Gärtner W., Jaeger K. E. and Eggert T.
Phys. Chem. Chem. Phys. (2005), 7: 2804-2811

Reproduced by permission of the PCCP Owner Societies

3.2 Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions.

Buttani V., Losi A., Eggert T., Krauss U., Jaeger K. E., Cao Z. and Gärtner W.
Photochem. Photobiol. Sci. (2007), 6: 41-49.

Reproduced by permission of The Royal Society of Chemistry (RSC) on behalf of the European Society for Photobiology and the European Photochemistry Association

3.3 Blue-light photoreceptors of the LOV-family in *Pseudomonas putida*: a conserved family sensor modules in saprotrophic Pseudomonads.

Manuscript in Preparation

RESEARCH PAPER

PCCP
www.rsc.org/pccp

Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct†

U. Krauss,[a] A. Losi,[b] W. Gärtner,[c] K.-E. Jaeger[a] and T. Eggert*[a]

[a] *Institut für Molekulare Enzymtechnologie, Heinrich-Heine Universität Düsseldorf, Forschungszentrum Jülich, D-52426 Jülich, Germany. E-mail: t.eggert@fz-juelich.de*
[b] *Dept. of Physics, University of Parma and CNR-INFM, Parma, Italy*
[c] *Max-Planck-Institut für Bioanorganische Chemie, Stiftstr. 34-36, 45470 Mülheim, Germany*

Received 1st April 2005, Accepted 3rd June 2005
First published as an Advance Article on the web 17th June 2005

The open reading frame PP2739 from *Pseudomonas putida* KT2440 encodes a 151 amino acid protein with sequence similarity to the LOV domains of the blue-light sensitive protein YtvA from *Bacillus subtilis* and to the phototropins (phot) from plants. This sensory box LOV protein, PpSB2-LOV, comprises a LOV core, followed by a C-terminal segment predicted to form an α-helix, thus constituting a naturally occurring paradigm for an extended LOV construct. The recombinant PpSB2-LOV shows a photochemistry very similar to that of YtvA and phot-LOV domains, yet the lifetime for the recovery dark reaction, τ_{rec} = 114 s at 20 °C, resembles that of phot-LOV domains (5–300 s) and is much faster than that of YtvA or YtvA-LOV (>3000 s). Time-resolved optoacoustics reveals phot-like, light-driven reactions on the ns–μs time window with the sub-nanosecond formation of a flavin triplet state (Φ_T = 0.46) that decays into the flavin–cysteine photoadduct with 2 μs lifetime (Φ_{390} = 0.42). The fluorescence spectrum and lifetime of the conserved W97 resembles the corresponding W103 in full-length YtvA, although the quantum yield, Φ_F, is smaller (about 55% of YtvA) due to an enhanced static quenching efficiency. The anisotropy of W97 is the same as for W103 in YtvA (0.1), and considerably larger than the value of 0.06, found for W103 in YtvA-LOV. Different to YtvA and YtvA-LOV, the fluorescence for W97 becomes larger upon photoproduct formation. These data indicate that W97 is located in a similar environment as W103 in full-length YtvA, but undergoes larger light-driven changes. It is concluded that the protein segment located C-terminally to the LOV core (analogous to an interdomain linker) is enough to confer to the conserved tryptophan the fluorescence characteristics typical of full-length YtvA. The larger changes experienced by W97 upon light activation may reflect a larger conformational freedom of this protein segment in the absence of a second domain.

Introduction

The *Pseudomonas putida* KT2440 gene PP2739 encodes a 151 amino acid (aa) protein (Q88JB0)[1] that comprises a LOV (light, oxygen and voltage) domain similar to the photosensing units of plant phototropins (phot),[2] and of the bacterial YtvA protein from *Bacillus subtilis*[3] LOV domains belong to the PAS (PerArntSim) superfamily, small protein modules of about 110 amino acids with a characteristic α–β folding, implied in a variety of sensory functions.[4] The LOV domains of YtvA and phot bind oxidized flavin mono-nucleotide (FMN) as a chromophore and absorb maximally at about 450 nm (LOV_{447}) in the dark state.[3,5,6] Blue-light illumination triggers a photocycle involving the reversible formation of a blue-shifted FMN-cysteine C(4a)-thiol adduct (LOV_{390}).[7] The lifetime for the dark recovery process, τ_{rec}, is between 5 and 330 s (25 °C) for phot-LOV domains,[8,9] whereas for YtvA and YtvA-LOV τ_{rec} is exceptionally long (2600 and 3900 s, at 25 °C).[10]

The *P. putida* LOV protein characterized in this work is referred to as PpSB2-LOV (SB2 = Sensory Box 2, the suffix 2 was taken in accordance with reference[1]). Besides the LOV domain, it contains an 18 aa long N-terminal sequence and a C-terminal extension of 31 aa, without a fused effector domain. A similar architecture is found in the photoactive yellow protein (PYP), where the light sensing unit is a PAS domain binding an isomerizable chromophore, with a folding similar to a LOV domain.[11]

In the LOV core of PpSB2-LOV all the residues known to interact with FMN in LOV domains are conserved with LOV1 or LOV2, with the exception of an asparagine in the canonical G*X*NCRFLQG sequence (*X* = any aa) that contains the active cysteine. This sequence is changed into YQ**D**CRFLQG in PpLOV-SB2, and is characteristic of a few other bacterial LOV proteins.[1] Such modification in the binding motif, *i.e.*, the introduction of a charged aa at the position preceding the reactive cysteine, raises questions about the effects of this exchange on the suggested photochemistry.

Furthermore, the additional 31 aa at the C-terminus of the PpSB2-LOV protein renders it a naturally occurring extended LOV protein without an effector domain, a structure that can be compared to a synthetic construct of a plant LOV domain, recently characterized by NMR spectroscopy.[12] In that work, the authors characterized the light-dependent conformational changes of an *Avena sativa* phot1-LOV2 extended construct comprising 40 aa downstream the LOV core (LOV2-40). Their results showed the formation of an about 20 aa long amphipathic helix (Jα) in the C-terminal extension that interacts with the LOV domain in the dark and undergoes conformational changes upon blue-light irradiation.[12] The NMR studies revealed also that the interaction of the single, conserved LOV

† Electronic supplementary information (ESI) available: Fig. 2 in colour; colour also in HTML version. See http://dx.doi.org/10.1039/b504554a

DOI: 10.1039/b504554a

tryptophan (W491) to the amphipatic helix is abolished upon light activation. Recent studies also determined that the conserved W of YtvA-LOV (W103) is implied in interdomain interaction, although it undergoes very small conformational changes in full-length YtvA.[13] The investigation of PpSB2-LOV, comparable to an extended LOV construct, may give additional information on the role of the conserved W in LOV domains.

In this work, we demonstrate that the light-driven reactions occurring in the PpSB2-LOV protein are similar to those in phot-LOV domains, despite the differences in the conserved sequence that contain the reactive cysteine. Time-resolved optoacoustic spectroscopy has been employed to investigate the early kinetics of the photocycle and thermodynamics of the transient species, together with determination of the photocycle quantum yields. UV fluorescence spectroscopy has been applied to the conserved W97 with the aim to understand whether the short C-terminal sequence is sufficient to mimic the properties of the homologous W103 in full-length YtvA. Furthermore, we wanted to understand whether the absence of a second domain (STAS in YtvA) renders the C-terminal segment to a more flexible conformation and responsive to the light-induced reactions, in order to compare the results with the NMR derived conclusions.

Materials and methods

Protein expression and purification

The Sensory Box 2 protein of *Pseudomonas putida* KT2440 (PpSB2-LOV) (SwissProt: Q88JB0) was PCR-amplified from genomic DNA of *P. putida* KT2440 using the primer pair SB2_NdeI_fw: 5′-CCT ACA ATC ATA TGA TCA ACG CAA AAC TCC TGC-3′, and SB2_XhoI_rev_Stop: 5′-CGC TAA CTC GAG TCA GTG CTT GGC CTG G-3′, and subsequently cloned into the pET28a plasmid (Novagen-Merck, Darmstadt, Germany), allowing the expression as N-terminal 6xHis-Tag fusion protein. The His-tagged PpSB2-LOV protein was expressed in *Escherichia coli* BL21 (DE3), induced by addition of 0.4 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Roth, Karlsruhe, Germany). Following sonication of the cells, the soluble PpSB2-LOV protein was purified from crude cell free extracts by affinity chromatography under native conditions on a Ni-NTA Superflow resin (QIAGEN, Hilden, Germany) using an ÄKTA-explorer™ chromatography system (Amersham Biosciences Ltd, Buckinghamshire, UK). The molecular weight of the purified PpSB2LOV fusion protein was estimated by size exclusion chromatography on an ÄKTA-explorer™ chromatography system using a HiLoad Superdex75 XK16/60 column (Amersham Biosciences Ltd, Buckinghamshire, UK) with 20 mM Tris/HCl pH 8.0, 0.15 M NaCl as eluent. Calibration was performed with a Low Molecular Weight Gel Filtration Calibration Kit (Amersham Bioscience). Full-length YtvA and YtvA-LOV domain (aa 26-127 of the full protein) were prepared as previously described.[10] *Chlamydomonas reinhardtii* phot-LOV1 (Cr-phot-LOV1) was kindly donated by Prof. Peter Hegemann (Humboldt University Berlin, Germany).

Modelling of three-dimensional structure

The three-dimensional structure of PpSB2-LOV was modelled on homology to phot-LOV1 domain of *C. reinhardtii*, obtained from the Brookhaven Protein Data Bank (1n9lA) and aligned using the tools available in the SwissPdbViewer.[14] The addition of water molecules to structural models was performed using the VegaZZ package.[15] Energy minimization was carried out with the MAB all-atom force field implemented in the MOLOC molecular modelling program suite[16,17] and the Gromos96 force field implementation of SwissPdbViewer.[18]

Preliminary docking of the C-terminal extension to the PpSB2-LOV-core

Secondary structure prediction for the 31 aa C-terminal extension of the PpSB2-LOV core (residues 122-152) was carried out using the PredictProtein server's PROF algorithm (http://www.predictprotein.org/).[19] Docking of the predicted helical structure segment to the protein core was done using the ClusPro Server (http://nrc.bu.edu/cluster/). The ClusPro docking algorithm evaluates billions of possible complexes, retaining a preset number with favourable surface complementarities. A filtering method is then applied to this set of structures, selecting those with good electrostatic and desolvation free energies for further clustering.[20] The program's output is a short list of possible complexes ranked according to their clustering properties. Evaluation of the model's quality was done using the SAVS (structure analysis and verification) server at http://www.doe-mbi.ucla.edu/, which employs the tools WHAT_CHECK,[21] PROCHECK,[22] ERRAT,[23] and VERIFY3D.[24]

Electrostatic calculations

Electrostatic calculations on the PpSB2-LOV core and the adjacent helical extension as well as the generation of the corresponding molecular electrostatic potential surface maps were performed using the VegaZZ package.[15]

Spectrophotometric analysis

Absorption spectra and kinetics curves were recorded using a Shimadzu UV-2102 PC spectrophotometer. Fluorescence measurements were carried out at 10 °C, in phosphate buffer 10 mM, pH = 8, NaCl 10 mM. FMN (FLUKA, Neu-Ulm, Germany) was used as a standard ($\Phi_F = 0.26$)[25] to measure the fluorescence quantum yield of the PpSB2-LOV flavin chromophore. Fluorescence lifetimes were measured with a time correlated single photon counting nFL920 spectrometer (Edinburgh instruments, Livingston, UK), equipped with M300 excitation and emission monochromators, a start PMT detector and a red light-sensitive S300-R PMT detector. Excitation at 280 and 295 nm was achieved with an F900 nanosecond flash-lamp (Edinburgh instruments, Livingston, UK), with hydrogen as operating gas. Pulse frequency was 40 kHz. Steady state fluorescence spectra were recorded with a Cary Eclipse fluorescence spectrophotometer (Varian, Victoria, Australia). Photoequilibrium conditions, with accumulation of the photoactivated state were achieved illuminating the sample with a blue-light emitting Led-Lenser®V8 lamp (Zweibrüder Optoelectronics, Solingen, Germany). Steady state anisotropy, r, has been measured with a Perkin-Elmer LS-50 luminescence spectrometer equipped with the polarization accessory Par No. L225 0100 (Perkin-Elmer Ltd., Beaconsfield, England). The anisotropy is defined by intensities of vertically ($I_{\parallel}$) and horizontally ($I_{\perp}$) polarized light measured when excitation light is vertically polarized, namely $r = (I_{\parallel} - I_{\perp})/(I_{\parallel} + 2I_{\perp})$.[26]

For the laser induced optoacoustics spectroscopy (LIOAS) experiments, excitation at 450 nm was achieved by pumping the frequency-tripled pulse of a Nd:YAG laser (SL 456G, 6 ns pulse duration, 355 nm, Spectron Laser System, Rugby, UK) into a β-barium borate optical parametric oscillator (OPO-C-355, bandwidth 420–515 nm, Laser Technik Vertriebs GmbH, Ertestadt-Friesheim, Germany) as previously described.[27] The cuvette holder FLASH 100 (Quantum Northwest, Spokane, WA, USA) was temperature controlled to ±0.02 °C. The signal was detected by a V103-RM ultrasonic transducer and fed into a 5662 preamplifier (Panametrics Inc., Waltham, MA, USA). The pulse fluence was varied with a neutral density filter and measured with a pyroelectric energy meter (RJP735 head connected to a meter RJ7620 from Laser Precision Corp.).

The beam was shaped by a 1 × 12 mm slit, allowing a time resolution at most of ~60 ns by using deconvolution techniques.[28] The experiments were performed in the linear regime of amplitude *versus* laser fluence and the total incident energy normally used was ~20 μJ pulse^{-1} (this corresponds to 7.5 × 10^{-11} einstein for 450 nm excitation, photon energy 265.82 kJ mol^{-1}). The sample concentration was about 15 μM, corresponding to 1.8 × 10^{-9} mol in the excitation volume $V_0 = 0.12$ mL. These conditions correspond to 0.04 photon per protein molecule. New coccine (FLUKA, Neu-Ulm, Germany) was used as calorimetric reference.[29] The time evolution of the pressure wave was assumed to be a sum of monoexponential functions. The deconvolution analysis yielded the fractional amplitudes (φ_i) and the lifetimes (τ_i) of the transients (Sound Analysis 3000, Quantum Northwest Inc., Spokane, WA, USA). The time window was between 20 ns and 5 μs. At a given temperature and for each resolved *i*-th step the fractional amplitude φ_i is the sum of the fraction of absorbed energy released as heat (α_i) and the structural volume change *per* absorbed einstein (ΔV_i), according to eqn. (1):[30,31]

$$\varphi_i = \alpha_i + \frac{\Delta V_i}{E_\lambda}\frac{c_p \rho}{\beta} \quad (1)$$

E_λ is the molar excitation energy, $\beta = (\partial V/\partial T)_p/V$ is the volume expansion coefficient, c_p is the heat capacity at constant pressure, and ρ is the mass density of the solvent. In this work we used the so-called "two temperature" (*TT*) method in order to separate α_i from ΔV_i.[32] The sample waveform was acquired at a temperature for which heat transport is zero, $T_{\beta=-0} = 3.2$ °C, and at a slightly higher temperature, $T_{\beta>0} = 10$ °C. At $T_{\beta=0}$ the LIOAS signal is only due to ΔV_i. The reference for deconvolution was recorded at $T_{\beta>0}$, and eqn. (2a) and (2b) were then used to derive α_i and ΔV_i:

$$\Delta V_i = \varphi_i \left| T_{\beta=0} E_\lambda \frac{\beta}{C_p \rho} \right|_{T_{\beta>0}} \quad (2a)$$

$$\alpha_i = \varphi_i \left| T_{\beta=0} - \varphi_i \right|_{T_{\beta=0}} \quad (2b)$$

The activation energy (E_a) for the recovery reaction was determined with the Arrhenius eqn. (3):

$$\ln k = -\left(\frac{E_a}{RT}\right) + \ln A \quad (3)$$

where $k = 1/\tau_{rec}$. Given that $E_a \approx \Delta H^\ddagger$ and $\Delta S^\ddagger \approx R \ln(Ah/k_B T)$, the free energy of activation, $\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$, can be estimated at a given temperature. The temperature range investigated was between 10 and 25 °C.

Results

Heterologous expression of PpSB2-LOV protein and purification

The PpSB2-LOV protein was heterologously expressed in *E. coli* BL21 (DE3) as an N-terminal His–Tag fusion protein. The protein had an apparent size of approximately 18 kDa as determined by SDS-PAGE analysis, which was in accordance with the calculated molecular weight for the fusion-protein (19.2 kDa). However, the size of the native protein as determined by analytical gel filtration experiments was about 39 kDa, which is two times larger than its calculated size, suggesting a dimeric organization of the native protein.

The soluble protein was purified chromatographically to electrophoretic homogeneity using immobilized metal ion affinity chromatography (IMAC). The eluted PpSB2-LOV fusion protein showed a yellow color, indicating the presence of an associated chromophore. Under blue light illumination (450 nm) the purified protein preparation emits a green–yellow fluorescence, which shows light-induced reversible bleaching.

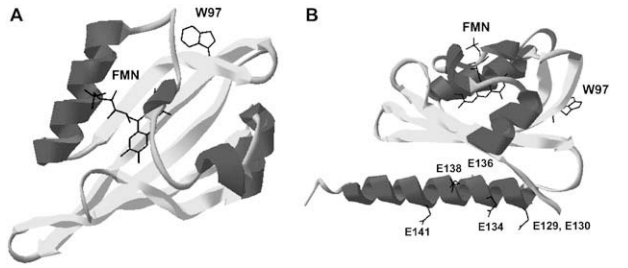


Fig. 1 (A) Theoretical model of the PpSB2-LOV protein LOV core, showing a typical LOV domain fold consisting of three α-helices which form together with a five-stranded β-scaffold the pocket around the FMN chromophore. (B) Theoretical model of the extended LOV core of PpSB2-LOV prepared using an automatic protein–protein docking approach. The C-terminal extension predicted to have an helical structure could be attached to the LOV-core of PpSB2-LOV. A cluster of negatively charged glutamic acid residues (E134, E136, E138, E141) in the LOV core's helical extending segment is highlighted. Furthermore both pictures show the conserved tryptophan (W97) residue and the chromophore flavin mononucleotide (FMN). Both figures were generated using the SwissPdbViewer.[14]

Three-dimensional structure modelling

Homology modelling of the PpSB2-LOV core results in a characteristic LOV domain folding, including the FMN chromophore (Fig. 1A) as described in detail for the *B. subtilis* phototropin-related protein YtvA.[3] However, in case of PpSB2-LOV the protein contains a C-terminal extension (residues 122–152) which shows α-helical structure in secondary structure predictions. To connect the C-terminal α-helical extension to the PpSB2-LOV core, an automatic docking approach was used. Therefore, a pdb-file containing the coordinates of the helical extension was produced using SwissPdbViewer which has been docked after energy minimization in an automatic manner to the PpSB2-LOV core using the ClusPro server.[20] The docking approach comprised a predicted structural complex in which the α-helical segment could be aligned to the PpSB2-LOV core's C-terminal end, spanning the five stranded β-scaffold surface of the protein core, placing the α-helical extension directly underneath the FMN chromophore. The two segments were joined using the tools available in the MOLOC molecular modelling program suite. Energy minimization was carried out as described in the materials and methods section. Fig. 1B shows the derived model, containing both the PpSB2-LOV core and the docked α-helical extension. According to electrostatic calculations and helical wheel analysis the C-terminal α-helix has amphipathic character containing a cluster of negatively charged glutamic acid residues forming a hydrophilic (E129, E130, E134, E136, E138, E141) and a nonpolar hydrophobic face (A128, V132, L135, V139, L142). The hydrophilic, negatively charged side of the helical segment probably faces the aqueous layer whereas the hydrophobic amino acids of the α-helical extension interact with the hydrophobic surfaces of the β-sheet of the LOV domain (Fig. 2) as shown in case of the extended LOV2-40 construct.[12]

Spectroscopical studies and recovery kinetics

The absorption spectrum of PpSB2-LOV$_{450}$ (dark state) is shown in Fig. 3, compared to the absorption spectrum of YtvA-LOV$_{450}$. Although the two spectra are very similar in the blue-light region around 350 nm (UVA), PpSB2-LOV is more similar to *Chlamydomonas reinhardtii* LOV1 (Cr-LOV1),[33] with the band being slightly blue-shifted and broader with respect to YtvA-LOV. This feature is reminiscent of the spectral differences between phot LOV1 and LOV2. The fluorescence spectrum of the bound chromophore, upon 450 nm excitation, is the same as in YtvA and LOV domains.[10] The

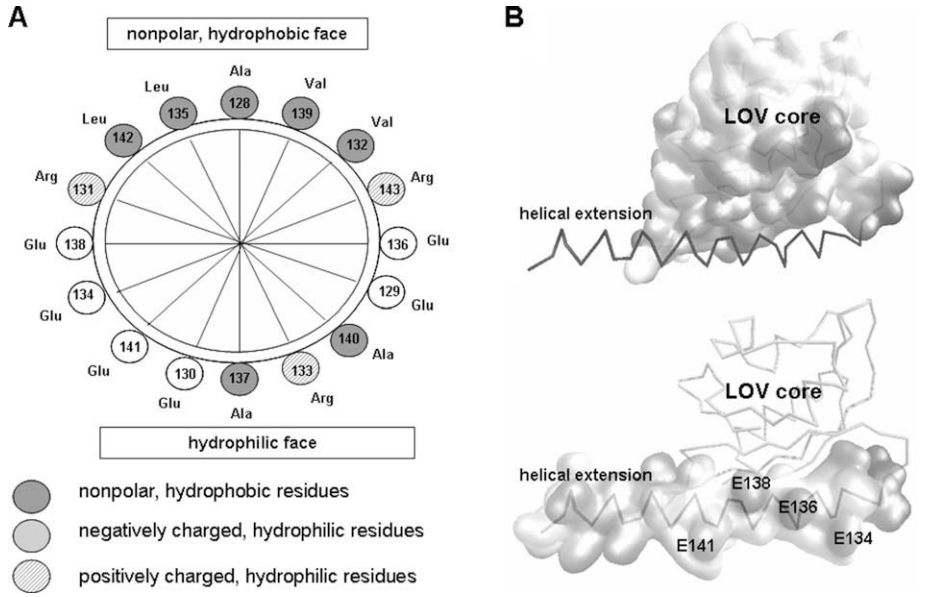


Fig. 2 (A) Helical wheel analysis of the α-helical part (residues 125–140) of the PpSB2-LOV C-terminal extension. (B) Electrostatic surface potential of the LOV core and the adjacent C-terminal helical extension with negatively charged regions shown in red and regions with positive charge in blue. Figure (B) was generated using the VegaZZ tool.[15]

steady state anisotropy of the bound flavin (for 450 nm excitation, emission recorded at 500 nm) is 0.30 ± 0.02 in all three proteins, indicating that the chromophore pocket is similarly rigid.

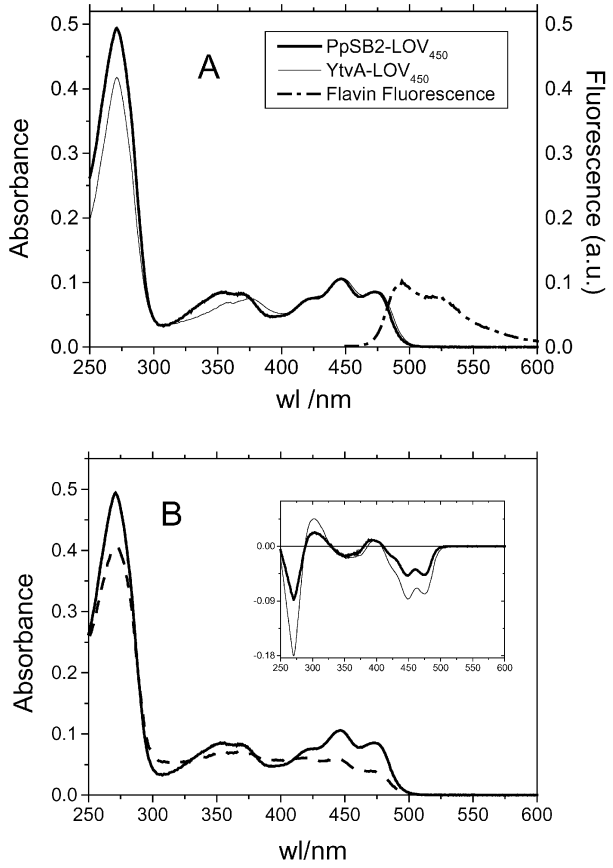


Fig. 3 (A): Absorption spectrum of the PpSB2-LOV protein in the dark (thick line) compared to YtvA-LOV (thin line). The fluorescence spectrum of the bound flavin chromophore after 450 nm excitation is also shown for PpSB2-LOV (dashed line). The intensity has been arbitrarily normalized in order to be compared with the absorbance. (B): changes in absorbance induced by blue-light illumination in PpSB2-LOV (dashed line). The inset shows the light–dark difference spectra for PpSB2-LOV (thick line) and YtvA-LOV (thin line).

Upon illumination with blue-light, the absorption band in the visible region decreases with concomitant loss of the flavin fluorescence, although the photoproduct of PpSB2-LOV cannot be accumulated to 100% like in YtvA and YtvA-LOV.[10] The difference spectrum shows that light activation induces the formation of a photoadduct, referred to as PpSB2-LOV$_{390}$. The recovery of the absorption in the dark can be followed at 480 nm,[10] and obeys a single exponential decay. The recovery lifetime, τ_{rec} = 114 s at 20 °C, is much faster than in the *B. subtilis* YtvA (τ_{rec} = 3200 s) and YtvA-LOV (τ_{rec} = 5280 s), and is similar to Cr-phot-LOV1 (τ_{rec} = 334 s) at the same temperature. τ_{rec} is strongly temperature dependent, showing a linear Arrhenius behaviour in the temperature range between 10 and 25 °C (see materials and methods). From the values of the activation energy E_a = 72.6 kJ mol^{-1} and the preexponential factor $A = 4.54 \times 10^{10}$ s^{-1}, the values of activation enthalpy ($\Delta H^{\ddagger}$), entropy ($\Delta S^{\ddagger}$), and free energy ($\Delta G^{\ddagger}$) can be derived (summarized in Table 1).

Comparison with data obtained for YtvA, YtvA-LOV and Cr-phot-LOV1 shows that the value of $\Delta G^{\ddagger}$ accounts for the differences in the observed τ_{rec}, *i.e.* higher $\Delta G^{\ddagger}$ correspond to slower kinetics. We note that the temperature range employed is very small, giving rise to larger errors than the experimental ones (within 10% from a set of two experiments for each sample) in the determination of the activation parameters. This is especially true for the value of $\Delta S^{\ddagger}$, derived from the intercept of the Arrhenius plots. Therefore the apparent correlation between τ_{rec} and $\Delta G^{\ddagger}$ must be taken with great care, given the large influence of the $T\Delta S^{\ddagger}$ term. The value of E_a is more reliable than $\Delta G^{\ddagger}$ and gives a correlation with τ_{rec} only for three out of the four proteins investigated, excluding YtvA-LOV. A larger array of LOV proteins would be needed in order to assess if a correlation holds for the majority of them.

Time-resolved optoacoustic (LIOAS) studies

The LIOAS signals at $T_{\beta=0}$ = 3.2 °C are similar to those recorded for YtvA, YtvA-LOV and Cr-phot-LOV1,[3,10,33] with a transient species formed within 20 ns with a small contraction ($\Delta V_1 = -0.8$ ml einstein^{-1}), decaying with a lifetime $\tau_2 = 2$ μs into a second transient that is formed with a larger volume contraction, $\Delta V_2 = -6$ ml/einstein (Table 2, Fig. 4).

In analogy with the other known LOV proteins, the unresolved step ($\tau_1 < 20$ ns) is assigned to the fast reactions

Table 1 Activation parameters for the recovery reaction in LOV proteins

	P. putida-SB2	[a]YtvA	[a]YtvA-LOV	[a]Cr-phot-LOV1
E_a/kJ mol^{-1}	72.6	98	68	81
A/s^{-1}	4.54×10^{10}	4.6×10^{13}	1.8×10^{8}	7.5×10^{11}
[c]$\Delta S^{\ddagger}$/J mol^{-1} K^{-1}	−40.7	16.65	−86.9	−17.6
$\Delta H^{\ddagger}$/kJ mol^{-1}	72.6	98.0	68.0	81.1
[c]$\Delta G^{\ddagger}$/kJ mol^{-1}	84.5	93.0	93.9	86.3
τ_{rec}/s (at 20 °C)	114	3200	5280	334

[a] Ref. 10. [b] Losi and Ternelli, unpublished data (*Cr*: *Chlamydomas reinhardtii*). [c] At 20 °C. The values are derived from a set of two measurements for each sample. The experimental error associated with E_a, A and τ_{rec} is about 10%.

Table 2 Raw LIOAS data for PpSB2-LOV as compared with other LOV proteins

	α_1 ($\tau_1 < 20$ ns)	ΔV_1/mL mol^{-1}	α_2 (τ_2 *ca.* 2 μs)	ΔV_2/mL mol^{-1}
PpSB2-LOV	0.43 ± 0.03	−0.8 ± 0.04	0.14 ± 0.05	−6.0 ± 0.6
[a]YtvA	0.3 ± 0.03	−0.44 ± 0.05	0.23 ± 0.04	−5.8 ± 0.2
[a]YtvA-LOV	0.33 ± 0.1	−0.46 ± 0.05	0.28 ± 0.1	−9.1 ± 0.7
[b]Cr-phot-LOV1	0.36 ± 0.04	−0.72 ± 0.1	0.09 ± 0.10 ($\tau_2 = 1.2$ μs)	−4.1 ± 0.8

[a] Ref. 10. [b] Ref. 33.

generating the FMN triplet state (referred to as PpSB2-LOV$_T$). The microsecond process ($\tau_2 = 2$ μs) corresponds to the decay of PpSB2-LOV$_T$ into the photoadduct, PpSB2-LOV$_{390}$.[3,10,33] With the knowledge that the FMN triplet state, both in solution and in LOV proteins lies at *ca.* 200 kJ mol^{-1},[3,33] it is possible to calculate the formation quantum yield of PpSB2-LOV$_T$, Φ_T, from eqn. (4):

$$\Phi_T \frac{E_T}{E_\lambda} = 1 - \alpha_1 - \Phi_F \frac{E_F}{E_\lambda} \quad (4)$$

where E_F is the average energy for the fluorescence emission (232 kJ mol^{-1}, 515 nm), $\Phi_F = 0.25$ is the flavin fluorescence quantum yield (measured in comparison to FMN and YtvA), and $E_\lambda = 265.8$ kJ mol^{-1} is the photon energy at 450 nm. The resulting value is $\Phi_T = 0.46 \pm 0.04$ (Table 3).

The value of $E_T = 200$ kJ mol^{-1} for the triplet state of the bound chromophore, as measured by means of LIOAS in LOV proteins, has been recently confirmed by phoshorescence spectroscopy on a phot-LOV2 domain,[34] that is intrinsically very precise. Therefore the error associated with Φ_T does not exceed the experimental one of about 10%.

The energy level of PpSB2-LOV$_{390}$ is calculated by means of eqn. (5) and requires the determination of Φ_{390}. This was done

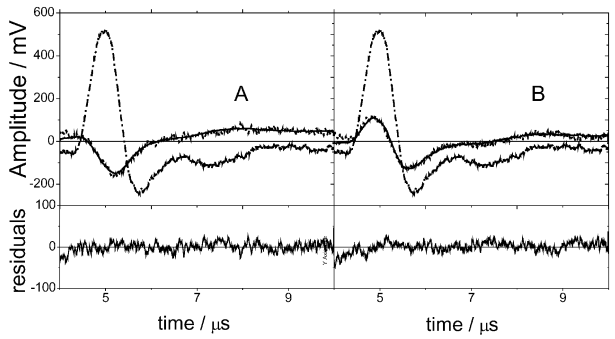


Fig. 4 (A) PpSB2-LOV LIOAS signal recorded at $T_{\beta=0} = 3.2$ °C (full line) and at $T_{\beta>0} = 10$ °C for the reference (new coccine, dashed line). At this temperature, the signal for the reference is zero, whereas the protein signal reflects structural volume changes (here contractions). The curve fitted to the experimental signal with a sum of two exponential functions (see text) is superimposed (dotted line) on the sample waveform. (B) Reference and sample recorded at 10 °C. At this temperature both heat deposition and structural volume changes contribute to the protein signal.

relatively to YtvA as previously reported, 10 giving $\Phi_{390} = 0.42 \pm 0.08$, namely close to 100% efficiency from the triplet state (the error derives from the uncertainty of the measurements and the one associated with Φ_{390} for YtvA).

$$\alpha_2 = \Phi_T \frac{E_T}{E_\lambda} - \Phi_{390} \frac{E_{390}}{E_\lambda} \quad (5)$$

Finally, the molecular volume changes are evaluated by means of eqn. (6a) and (6b):

$$\Delta V_T = \frac{\Delta V_1}{\Phi_T} \quad (6a)$$

$$\Delta V_{390} = \Delta V_T + \frac{\Delta V_2}{\Phi_{390}} \quad (6b)$$

The resulting values, photophysical, photochemical and structural parameters, are reported in Table 3, and are compared with other LOV proteins. The largest errors are associated with the determination of E_{390}, whose value, different than that for E_T, cannot be checked by means of optical techniques. Furthermore experimental errors are propagated through eqn. (5) and the determination of α_i with LIOAS has intrinsically a large uncertainty.[35] As a whole the uncertainty associated with E_{390} is $\geq 50\%$.

Fluorescence of the conserved W97

The PpSB2-LOV protein possesses a single tryptophan, W97, localized in the LOV core and conserved with the majority of LOV domains (Fig. 1).[1] The fluorescence maximum of W97 in PpSB2-LOV and YtvA is localized at about 336 nm, with a half-band width of about 60 nm, similarly to W103 in YtvA.[13] In YtvA-LOV, W103 has a maximum at about 345 nm and a broader half-band width (68 nm).[13] The steady state anisotropy of W97 (excitation at 295 nm, emission at 350 nm, 20 °C), $r = 0.1 \pm 0.02$ is also very similar to $r = 0.11 \pm 0.02$ for W103 in YtvA, and is higher than $r = 0.06 \pm 0.01$ in YtvA-LOV (the errors derive from a set of 5 measurements for each sample). This implies that W97 has a conformational flexibility similar to W103 in full-length YtvA. The fluorescence lifetime distribution is also comparable to W103 in YtvA (Table 4), although the relative fluorescence quantum yield of W97 (Φ_F) is only 55% in comparison to that of W103. This shows that the value of Φ_F is affected by ultrafast deactivation processes, *e.g.*

Table 3 Molecular light-driven volume changes and photochemical parameters

	[c]Φ_T	[d]E_{390}/kJ mol^{-1}	[e]Φ_{390}	[f]ΔV_T/mL mol^{-1}	[g]ΔV_{390}/mL mol^{-1}
PpSB2-LOV	0.46	133	0.42	−1.74	−16.0
[a]YtvA	0.62	136	0.49	−0.71	−12.5
[a]YtvA-LOV	0.69	113	0.55	−0.67	−17.2
[b]Cr-phot-LOV1	0.63	171	0.6	−1.15	−8.0

[a] Ref. 10. [b] Ref. 33. [c] Error is about 10%. [d] Error $\geq$ 50%; [e] Error is about 20%; [f] Error is about 10%; [g] Error is about 20%.

Table 4 Time-resolved fluorescence of W97 compared to W103 in YtvA and YtvA-LOV

	λ_{ex} = 295 nm, λ_{em} = 340 nm, T = 10 °C							
	τ_1/ns	A_1 (%)	τ_2/ns	A_2 (%)	τ_3/ns	A_3(%)	$\langle\tau\rangle$/ns	χ^2
PpSB2-LOV			1.8	37	5.8	63	4.3	1.055
YtvA			2	24	5.6	76	4.7	1.101
YtvA-LOV	0.87	33	2.5	46	5.8	21	2.6	1.034

The errors associated with lifetimes (τ_i) and relative amplitudes (A_i) are within 5%. Average of two sets of experiments for each sample.

formation of exciplexes with neighbouring polar groups and/or other static quenching phenomena.

Different to W103 in YtvA and YtvA-LOV, the fluorescence of W97 is affected by the formation of the photoproduct. It increases of about 7% in the photo-equilibrium state. This change is fully reverted in the dark with a $\tau_{recW} = 102 \pm 10$ s at 20 °C, quite in accordance to $\tau_{rec} = 114$ s for the recovery of the flavin absorption (Fig. 5). The difference in the fluorescence intensity between the dark and photoequilibrium states is, in absolute terms, very small, giving rise to a large noise in the kinetics traces and a quite large error in the value of τ_{recW}. The data are nevertheless fully reproducible and the light-induced effect completely reversible, showing that the difference is significative.

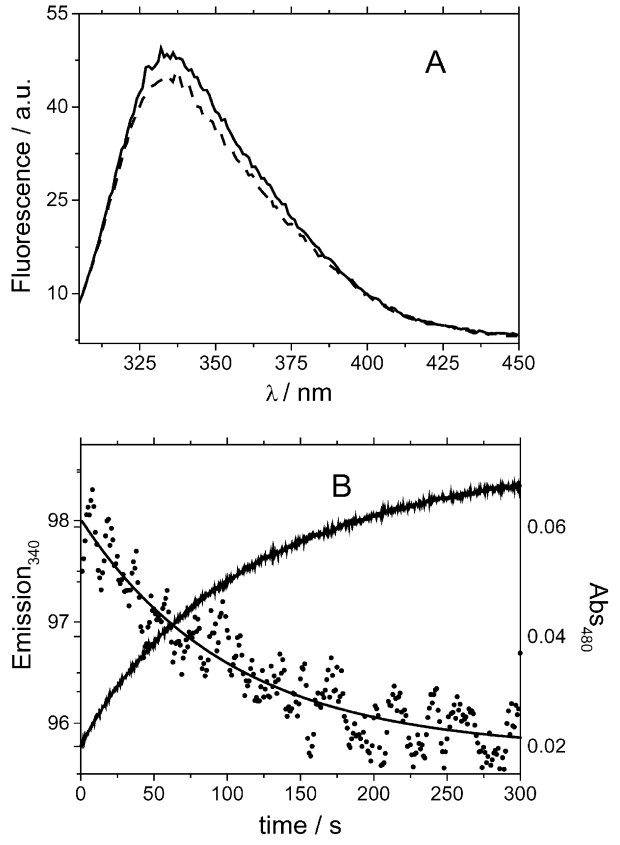


Fig. 5 (A) Fluorescence of W97 in PpSB2-LOV after 295 nm excitation in the dark (dashed line) and in the photo equilibrium state (about 50% photoproduct, full line). (B) Recovery kinetics for W97 in the dark (dots), followed by recording the fluorescence at 340 nm (excitation 295 nm). The monoexponential fitting of the experimental points (full line) yields $\tau_{rec} = 102 \pm 10$ s at 20 °C. For comparison the time course of absorption recovery of the flavin chromophore at 480 nm has been reported.

Discussion

In this work we have demonstrated that the PpSB2-LOV protein undergoes phot-like photochemistry, similarly to other prokaryotic proteins, *i.e.* YtvA from *B. subtilis*[3] and a LOV-kinase from *Caulobacter crescentus*.[1] PpSB2-LOV represents a novel variation of LOV proteins, in that the canonical G*X*NCRFLQG sequence is changed for YQ**D**CRFLQG, without impairing chromophore binding or the photochemical activity of the protein. Interestingly, the native PpSB2-LOV protein seems to have a dimeric organization as identified by size exclusion chromatography. Very recently, this dimerization effect has been suggested for other LOV-domain containing proteins by different groups: Nakasako and co-workers reported about the dimerization of the LOV-domain containing protein (FKF1) of *Arabidopsis thaliana*[36] and Salomon *et al.* discussed the possible LOV1 mediated dimerization of phototropin of *Avena sativa*.[37] The phototropin-like protein YtvA from *B. subtilis* also appears as dimer after overexpression in *E. coli* and purification by affinity chromatography, when treated under non-reducing conditions (Krauss and Eggert, unpublished results). However, it cannot be excluded that the observed dimer formation might be an artifact linked to the purification conditions. Its importance for the protein function *in vivo* still has to be determined.

As evidenced from the LIOAS data (Table 2), the thermodynamic profile and photocycle kinetics (Table 1) and quantum yield (Table 3) are comparable with previously reported values in other LOV proteins.[10,33] The photoproduct has a high energy content, $E_{390} = 130$ kJ mol^{-1}, indicating a strained conformation, typical of these blue-light sensors. In contrast, the energy level of the signalling state in PYP is only about 60 kJ mol^{-1}, suggesting a relaxed protein structure.[38]

The study of the fluorescence properties of the conserved W97 (Table 4, Fig. 5) also gives information on the molecular mechanisms underlying interdomain communication and light-to-signal conversion in LOV proteins. In phot the photochemical reactions centered at the LOV domains trigger self-phosphorylation through the C-terminal kinase moiety, *via* still poorly characterized inter-domain communication. In the bacterial LOV proteins, a single LOV can be associated to diverse output domains (*e.g.* kinases, phosphodiesterases,

response regulators) with putative enzymatic or regulatory activities.[3,39] In *B. subtilis* YtvA, the N-terminal LOV domain is followed by a STAS domain (Sulfate transporter/Anti-Sigma-factor antagonist), suggested to have a general NTP (nucleoside triphosphate) binding role.[40] It is still to be determined whether in the various LOV proteins the light-to-signal transduction mechanism is similar, even due to the lack of structural information on full proteins.

The recently characterized extended construct LOV2-40, possess a 20 aa long amphipathic helix (Jα), C-terminal to the LOV core and connected to it by a flexible loop of 10 aa.[12] In LOV2-40, the Jα helix affects the LOV domain structure in the dark, but the constraint is removed after light activation and the helix becomes disordered. This conformational change is reversible in the dark and the recovery can be time-resolved by following the change in the chemical shift of several amino acids involved in the interaction with the Jα helix,[41] including the conserved W491 of of LOV2-40. In contrast with these results, fluorescence studies on the corresponding W103 in YtvA showed that this residue undergoes very small changes upon formation of the photoproduct. Nevertheless, the fluorescence properties of W103 in YtvA are sharply different from those in the YtvA-LOV construct, indicating that W103 participates in inter-domain interaction. Similarly, W491 has a different chemical shift in LOV2-40 and in the LOV2 core, in the dark state.[12] The architecture of the PpSB2-LOV protein is similar to the extended LOV2-40 construct. The fluorescence results presented here (Table 4) show that the short C-terminal extension of PpSB2-LOV (31 aa) is sufficient to confer to W97 fluorescence properties (emission maximum, anisotropy and fluorescence lifetime distribution) similar to full-length YtvA having a second domain of 135 aa. This supports the idea that the conserved W participates in inter-domain communication by interacting with a region immediately following the LOV core, rather than *via* direct interaction with the effector domain. Contrary to the case of YtvA, but similar to LOV2-40, W97 in PpSB2-LOV undergoes well detectable and reversible changes in the photoequillibrium state (Fig. 5), suggesting a larger flexibility of this region in the absence of a second domain.

Secondary structure predictions indicate that the region C-terminal to the LOV core assumes a helical conformation in PpSB2-LOV, similar to the Jα helix that follows phot-LOV2. Furthermore, an automated docking approach of the C-terminal α-helical domain and the LOV core resulted in an extended structural model for PpSB2-LOV (Figs. 1 and 2) showing high similarities to the LOV2-40 construct reported by Harper *et al.*,[12] suggesting similar interactions between the LOV core and its adjacent helical extension.

The conserved W is located in the vicinity of the loop that connects the LOV core to the helical extension, both in our model of PpSB2-LOV and in the NMR-derived structure of LOV2-40,[12] appearing quite exposed to the solvent. This does not immediately explain why the fluorescence properties of W97 and W103 in YtvA are so different from those of YtvA-LOV, and why the chemical shift of W491 is sharply different between the extended LOV2-40 and the LOV2-core.[12] The reason may rely in the fact that the presence of the helical extension can affect the C-terminal part of the LOV core that appears to contain a looped region in the extended construct,[12] different from LOV domains of known structure, which are in β-strand.[42,43] This more flexible region could interact with W97 more strongly than in our model and extend to the loop that preceeds the helical extension.

The spectroscopic results presented in this work, together with the output of the docking approach, suggest that the mechanism of transmitting the light-induced conformational changes to the effector domains is similar between bacterial LOV proteins and plant phot. Furthermore, they highlight the importance of the linker region in determining the inter-domain interactions in this type of light sensors. Bacterial LOV proteins, with their variety of effector domains or even occurring as extended LOV-constructs[1] may thus represent good systems to understand the molecular details of signal transmission from a LOV-core to domain partners.

Abbreviations

PAS(PerArntSim): Per: period circadian protein, Arnt: Ah receptor nuclear translocator protein, Sim: single-minded protein; STAS: Sulfate Transporter/AntiSigma factor antagonist; LOV: Light, Oxygen and Voltage; PYP: photoactive yellow protein; W: tryptophan.

Acknowledgements

This work has been supported in part by the Deutsche Forschungsgemeinschaft (Forschergruppe FOR526). We thank Silvia E. Braslavsky for the use of the LIOAS facility and Elisa Ghiraldelli for the fluorescence measurements.

References

1 A. Losi, *Photochem. Photobiol. Sci.*, 2004, **3**, 566–574.
2 W. R. Briggs, J. M. Christie and M. Salomon, *Antioxid. Redox Signaling*, 2001, **3**, 775–788.
3 A. Losi, E. Polverini, B. Quest and W. Gärtner, *Biophys. J.*, 2002, **82**, 2627–2634.
4 B. L. Taylor and I. B. Zhulin, *Microbiol. Mol. Biol. Rev.*, 1999, **63**, 479–506.
5 J. M. Christie, M. Salomon, K. Nozue, M. Wada and W. R. Briggs, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 8779–8783.
6 W. R. Briggs, C. F. Beck, A. R. Cashmore, J. M. Christie, J. Hughes, J. A. Jarillo, T. Kagawa, H. Kanegae, E. Liscum, A. Nagatani, K. Okada, M. Salomon, W. Rüdiger, T. Sakai, M. Takano, M. Wada and J. C. Watson, *Plant Cell*, 2001, **13**, 993–997.
7 M. Salomon, J. M. Christie, E. Knieb, U. Lempert and W. R. Briggs, *Biochemistry*, 2000, **39**, 9401–9410.
8 M. Kasahara, T. E. Swartz, M. A. Olney, A. Onodera, N. Mochizuki, H. Fukuzawa, E. Asamizu, S. Tabata, H. Kanegae, M. Takano, J. M. Christie, A. Nagatani and W. R. Briggs, *Plant Physiol.*, 2002, **129**, 762–773.
9 S. Crosson and K. Moffat, *Plant Cell*, 2002, **14**, 1067–1075.
10 A. Losi, B. Quest and W. Gärtner, *Photochem. Photobiol. Sci.*, 2003, **2**, 759–766.
11 K. J. Hellingwerf, J. Hendriks and T. Gensch, *J. Phys. Chem. A*, 2003, **107**, 1082–1094.
12 S. M. Harper, L. C. Neil and K. H. Gardner, *Science*, 2003, **301**, 1541–1544.
13 A. Losi, E. Ternelli and W. Gärtner, *Photochem. Photobiol.*, 2004, **80**, 150–153.
14 N. Guex and M. C. Peitsch, *Electrophoresis*, 1997, **18**, 2714–2723.
15 A. Pedretti, L. Villa and G. Vistoli, *J. Mol. Graph. Model.*, 2002, **21**, 47–49.
16 P. R. Gerber and K. Muller, *J. Comput. Aid. Mol. Des.*, 1995, **9**, 251–268.
17 P. R. Gerber, *J. Comput. Aid. Mol. Des.*, 1998, **12**, 37–51.
18 W. R. P. Scott, P. H. Hunenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kruger and W. F. van Gunsteren, *J. Phys. Chem. A*, 1999, **103**, 3596–3607.
19 B. Rost, G. Yachdav and J. F. Liu, *Nucleic Acid Res.*, 2004, **32**, W321–W326.
20 S. R. Comeau, D. W. Gatchell, S. Vajda and C. J. Camacho, *Bioinformatics*, 2004, **20**, 45–50.
21 R. W. Hooft, G. Vriend, C. Sander and E. E. Abola, *Nature*, 1996, **381**, 272–272.
22 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornthon, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
23 C. Colovos and T. O. Yeates, *Protein Sci.*, 1993, **2**, 1511–1519.
24 R. Luthy, J. U. Bowie and D. Eisenberg, *Nature*, 1992, **356**, 83–85.
25 P. W. van den Berg, J. Widengren, M. A. Hink, R. Rigler and A. G. Visser, *Spectrochim. Acta A*, 2001, **57**, 2135–2144.
26 A. S. Ladokhin, in *Encyclopedia of Analytical Chemistry*, ed. R. A. Meyers, John Wiley & Sons Ltd, Chichester, 2000, pp. 5762–5779.
27 A. Losi, A. A. Wegener, M. Engelhard, W. Gärtner and S. E. Braslavsky, *Biophys. J.*, 2000, **78**, 2581–2589.
28 J. E. Rudzki, J. L. Goodman and K. S. Peters, *J. Am. Chem. Soc.*, 1985, **107**, 7849–7854.

29 S. Abbruzzetti, C. Viappiani, D. H. Murgida, R. Erra-Balsells and G. M. Bilmes, *Chem. Phys. Lett*., 1999, **304**, 167–172.
30 S. E. Braslavsky and G. E. Heibel, *Chem. Rev*., 1992, **92**, 1381–1410.
31 J. Rudzki-Small, L. J. Libertini and E. W. Small, *Biophys. Chem*., 1992, **41**, 29–48.
32 S. Malkin, M. S. Churio, S. Shochat and S. E. Braslavsky, *J. Photochem. Photobiol. B: Biol*., 1994, **23**, 79–85.
33 A. Losi, T. Kottke and P. Hegemann, *Biophys. J*., 2004, **86**, 1051–1060.
34 M. Gauden, S. Crosson, I. H. M. van Stokkum, R. van Grondelle, K. Moffat and J. T. M. Kennis, *Proc. SPIE*, 2004, **5463**, 97–104.
35 A. Losi and S. E. Braslavsky, *Phys. Chem. Chem. Phys*., 2003, **5**, 2739–2750.
36 M. Nakasako, D. Matsuoka, K. Zikihara and S. Tokutomi, *FEBS Lett*., 2005, **579**, 1067–1071.
37 M. Salomon, U. Lempert and W. Rüdiger, *FEBS Lett*., 2004, **572**, 8–10.
38 K. Takeshita, Y. Imamoto, M. Kataoka, F. Tokunaga and M. Terazima, *Biochemistry*, 2002, **41**, 3037–3048.
39 S. Crosson, S. Rajagopal and K. Moffat, *Biochemistry*, 2003, **42**, 2–10.
40 L. Aravind and E. V. Koonin, *Curr. Biol*., 2000, **10**, R53–R55.
41 S. M. Harper, L. C. Neil, I. J. Day, P. J. Hore and K. H. Gardner, *J. Am. Chem. Soc*., 2004, **126**, 3390–3391.
42 S. Crosson and K. Moffat, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 2995–3000.
43 R. Fedorov, I. Schlichting, E. Hartmann, T. Domratcheva, M. Fuhrmann and P. Hegemann, *Biophys. J*., 2003, **84**, 2492–2501.

Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV–LOV dimerization and interdomain interactions†

Valentina Buttani,[a] Aba Losi,*[a,b] Thorsten Eggert,[c] Ulrich Krauss,[c] Karl-Erich Jaeger,[c] Zhen Cao[d] and Wolfgang Gärtner[d]

Received 20th July 2006, Accepted 6th October 2006
First published as an Advance Article on the web 27th October 2006
DOI: 10.1039/b610375h

The Bacillus subtilis protein YtvA is related to plant phototropins in that it senses UVA–blue-light by means of the flavin binding LOV domain, linked to a nucleotide-binding STAS domain. The structural basis for interdomain interactions and functional regulation are not known. Here we report the conformational analysis of three YtvA constructs, by means of size exclusion chromatography, circular dichroism (CD) and molecular docking simulations. The isolated YtvA-LOV domain (YLOV, aa 25–126) has a strong tendency to dimerize, prevented in full-length YtvA, but still observed in YLOV carrying the N-terminal extension (N-YLOV, aa 1–126). The analysis of CD data shows that both the N-terminal cap and the linker region (aa 127–147) between the LOV and the STAS domain are helical and that the central β-scaffold is distorted in the LOV domains dimers. The involvement of the central β-scaffold in dimerization is supported by docking simulation of the YLOV dimer and the importance of this region is highlighted by light-induced conformational changes, emerging from the CD data analysis. In YtvA, the β-strand fraction is notably less distorted and distinct light-driven changes in the loops/turn fraction are detected. The data uncover a common surface for LOV–LOV and intraprotein interaction, involving the central β-scaffold, and offer hints to investigate the molecular basis of light-activation and regulation in LOV proteins.

Introduction

B. subtilis YtvA is a blue-light responsive protein (261 aa), carrying a flavin-binding LOV (light, oxygen, voltage) domain, with LOV belonging to the PAS (PerArntSim) superfamily.[1] The photochemical LOV paradigm has emerged during the last years, thanks to the discovery and molecular characterization of phototropins (phot), blue-light receptors for a variety of responses in plants.[2–4] Phot are organized in two N-terminal LOV domains (LOV1 and LOV2, *ca.* 110 amino acids) and a C-terminally located ser/thr kinase domain. A self-phosphorylation reaction[5,6] is activated by UVA–blue light illumination of phot, thanks to the activation the LOV domains that carry a flavin-mononucleotide (FMN) as chromophore.[7] Light activation of LOV domains triggers a photocycle that involves the reversible formation of a covalent adduct between a conserved cysteine residue and position C(4a) of FMN,[7–11] formed upon decay of the FMN triplet state.[12–14] Despite the similarities in the light-triggered reactions, the two phot-LOV domains have not the same functional significance, with the self-phosphorylation reaction being mostly mediated by LOV2,[15,16] that also presents a larger quantum yield for the formation of the covalent adduct.[9] The LOV paradigm is configuring as one of the most conserved among distant phyla, and LOV proteins are now well documented in eukaryotes and prokaryotes.[17,18] In bacteria they are present in about 15–17% of the sequenced genomes, and light-induced, phot-like reactions have been demonstrated for some of them.[18–20]

Although the mechanistic details of the photochemical FMN–Cys adduct formation are still under debate (see ref. 19 and references therein, and ref. 21) the photocycle of LOV domains is by far the best characterized part of the light-to-signal transduction chain, thanks to the fact that isolated LOV domains are readily expressed and have been functionally and structurally analyzed.[9,10,22,23] A structure of full-length phot is not yet available and little is known about the mechanism leading to kinase activation and on the protein surfaces involved in domain–domain interactions. Light-driven unfolding of the helical linker connecting LOV2 and the kinase domain (Jα-linker) has been recently proposed to trigger the self-phosphorylation reaction, based on NMR spectroscopy and mutagenesis experiments.[24,25] This idea has recently been challenged by the observation that the Jα-linker is not needed either for LOV2-kinase interaction or for light-driven phosphorylation of a heterologous substrate.[15]

In prokaryotes the LOV light-sensing module is coupled to diverse effector domains, such as kinases (similar to phot), phosphodiesterases, response regulators, DNA-binding transcription

[a]*Dept. of Physics, University of Parma, via G.P. Usberti 7/A, 43100-Parma, Italy. E-mail: losia@fis.unipr.it*
[b]*CNR-INFM, Parma, Italy*
[c]*Institut für Molekulare Enzymtechnologie, Heinrich-Heine Universität Düsseldorf, Forschungszentrum Jülich, D-52426 Jülich, Germany*
[d]*Max-Planck-Institut für Bioanorganische Chemie, Stiftstr. 34-36, 45470 Mülheim, Germany*

† Electronic supplementary information (ESI) available: Calculation of accessible and buried surface areas using the VADAR tool. See DOI: 10.1039/b610375h

factors, regulators of stress sigma factors.[17,18,26] Therefore, besides the intrinsic interest regarding their structure, function and physiological role, they also represent a powerful tool to understand fundamental and still open questions in the field of LOV-based photoperception. (i) Are the light-induced reactions, centered on the LOV domain, transmitted to effector partners by means of the same molecular mechanisms, and do LOV domains interact with partner domains by means of the same protein surface? In phot1-LOV2 the central β-scaffold has been demonstrated to participate in interdomain communication, making contact with the Jα-linker.[24,25] A similar process has been observed with phot2-LOV2.[27] (ii) Why only one LOV domain is present in bacterial LOV proteins, whereas phot possess two of such units organized in tandem?[17,18] The LOV2 domain has a higher photocycle quantum yield than LOV1,[9] and acts as the principal light-sensing domain triggering phot1 and phot2 kinase activity, whereas LOV1 might have a regulative role.[16] The amino acid sequence of bacterial LOV domains has in general intermediate characteristics between LOV1 and LOV2.[18] (iii) Which factors govern LOV–LOV dimerization, considered a key feature in PAS-mediated sensing/regulation,[28,29] and which is the relevance of it during light sensing? It was shown by gel filtration chromatography that phot1-LOV1 has a tendency to dimerize, whereas LOV2 is monomeric.[30] This has led to the suggestion that LOV1 is responsible for phot dimerization, providing a possible functional role for the tandem organization of LOV domains in phot.[30] By means of pulsed thermal grating Terazima and coworkers, detected a transient volume increase (about 1.8 times, with time constant of 300 μs) during light activation of an extended phot1-LOV2 construct (including an N-terminal cap and the Jα-linker), and interpreted this phenomenon as a transient dimerization,[31] whose functional significance is not known. Dimeric states have been detected by means of small-angle X-ray scattering (SAXS) for the LOV domain of FKF1[32] and phot LOV1 domains.[33] The SAXS experiments showed that phot1-LOV2 is a dimer (in contrast with ref. 30 and 31) whereas phot2-LOV2 is monomeric.[33] The LOV domain of WC-1 from *Neurospora* has also been shown to homodimerize *in vitro.*[34]

In this work we have investigated three different constructs of YtvA, in order to partially address these problems. In YtvA the LOV domain is linked to a C-terminal STAS domain (sulfate transporters antisigma-factor antagonists).[35] This architecture is conserved in LOV proteins form other Firmicutes, *e.g.* in *Listeria* and *Oceanobacillus* genera.[18] Recent work has shown that YtvA is a positive regulator in the environmental signaling pathway that activates the general stress factor σ^{B}[36,37] and, most importantly, that the cysteine involved in the photoadduct formation is needed for its *in vivo* function,[37] in turn regulated by blue-light activation.[38] These last two recent studies allow to regard Ytva as a real flavin-based blue-light photoreceptor in *B. subtilis*, not only a blue-light sensitive protein. The STAS is thought to be the effector domain of YtvA, although little is known of its molecular functionality, with the exception that it confers to YtvA the ability to bind GTP and ATP,[39] in analogy with another STAS protein.[40] The constructs that we used are the LOV core (YLOV, aa 25–126), the N-YLOV comprising also the first 24 aa (aa 1–126) and the full-length protein YtvA. We applied gel filtration chromatography to detect possible dimers and circular dichroism spectroscopy for secondary structure determination, improving the data analysis with respect to previous work.[41] The data uncover a common surface for YLOV homodimerization and interdomain interactions, and corroborate a molecular model of the YLOV dimer obtained by docking simulations. Similarities and differences with phot-LOV domains are discussed.

Experimental

Protein samples and chemicals

For the N-YLOV protein, the DNA sequence encoding LOV core + N-terminal cap (aa 1–126 of the full protein) was amplified by PCR (Polymerase Chain Reaction). The recombinant plasmid of full-length YtvA in (pET28a) was used as template, the primers were:

5′-CAGCCATATGGCTAGTTTTCAATCATT (forward)

5′-TATTACTCGAGTTAGGTGATATCATTCTGAATTC (reverse)

Platinum® Taq DNA Polymerase (Invitrogen, Karlsruhe, Germany) was used for the PCR. PCR product, digested with NdeI/XhoI (NEB, Ipswich, UK), was ligated into the expression vector pET28a (Novagen-Merck, Darmstadt, Germany), which was digested with the same restriction enzymes. An N-terminal extension, including the 6×His-tag (sequence: MGSSHHHH-HHSSGLVPRGSH) was furnished, in the same way as for YtvA and YLOV. For the details of full-length His-tagged YtvA and its isolated LOV core (YLOV) generation, see previous reports.[19,42] The His-tagged proteins were expressed in *E. coli* BL21 DE3 (Stratagene, Amsterdam, The Netherlands) using IPTG (BioMol, Hamburg, Germany) induction. The proteins were then purified by affinity chromatography on Talon (Qiagen, Hilden, Germany) and finally concentrated in Na-phosphate buffer 10 mM, NaCl 10 mM, pH = 8.

Chromatography

Gel filtration chromatography experiments were performed on a Pharmacia FPLC apparatus, using a Superdex 75 HR 10/30 column (Amersham Biosciences), equilibrated with Na-phosphate 10 mM, pH = 8, NaCl = 0.15 M. A calibration curve was made using bovine serum albumin (69 kDa), ovalbumin (42.7 kDa), α-chymotrypsin (25 kDa), myoglobin (16.9 kDa) and ribonuclease (13.7 kDa) (low M_w calibration kit, Amersham Biosciences). YtvA, YLOV and N-YLOV were loaded on the column at a concentration between 1 and 50 μM, to give a final concentration ranging from 0.05 to 2.5 μM at the detection peak, due to dilution through the column.

Circular dichroism spectroscopy and data analysis

Circular dichroism (CD) experiments were carried out using a Jasco J715 spectropolarimeter, calibrated with ammonium d-10-camporsulfonic acid. The measurements were carried out in the far-UV spectral region (195–240 nm) at a temperature of 20 °C and the buffer background was always subtracted. The optical pathlength was 0.2 cm. Protein concentration was estimated from the absorption coefficient at 220 nm, $\varepsilon_{220}^{YTVA} = 492800\ M^{-1}\ cm^{-1}$, $\varepsilon_{220}^{YLOV} = 223900\ M^{-1}\ cm^{-1}$ and $\varepsilon_{220}^{N-YLOV} = 267900\ M^{-1}\ cm^{-1}$, calculated by comparison with $\varepsilon_{447}^{FMN} = 12500\ M^{-1}\ cm^{-1}$, in a 1 : 1 protein to chromophore ratio (*vide infra*). The corresponding value for FMN is $\varepsilon_{220}^{FMN} = 34500\ M^{-1}\ cm^{-1}$, thus introducing a negligible error also in case that some apoprotein is present. The mean

residue ellipticity Θ_{MRW} was calculated from the concentration of residues, c, (281 aa for YtvA, 122 aa for YLOV and 146 aa for N-YLOV, including the tag, according to the formula $\Theta_{MRW} = \Theta_{obs}/(10 \times cl)$, where c is in mol liter^{-1}, $l = 0.2$ cm and Θ_{obs} is in mdeg and Θ_{MRW} in deg cm^2 dmol^{-1}. Typically, the protein concentration was in the μM range and $c = 10^{-6} \times 281 = 2.8 \times 10^{-4}$ mol liter^{-1} for YtvA, $c = 10^{-6} \times 122 = 1.2 \times 10^{-4}$ mol liter^{-1} for YLOV and $c = 10^{-6} \times 146 = 1.4 \times 10^{-4}$ mol liter^{-1} for N-YLOV. Prediction of secondary structure composition was performed using the convex constraint analysis (CCA) algorithm,[43,44] and an extended curves dataset comprising 46 protein spectra.[45] In CCA, the sum of the fractional weights of each component spectrum is constrained to be 1. In addition, a constraint-called volume minimization is defined which allows a finite number of component curves to be extracted from a set of spectra without relying on spectral nodes. CCA does not use X-ray crystallographic data in the deconvolution procedure. Once the basis curves are obtained, they must be assigned to specific secondary structures. The secondary structure was also predicted from the amino acid sequence by means of bioinformatic tools, using the consensus secondary structure prediction method at the Pôle Bioinformatique Lyonnais.[46]

Docking simulation, evaluation of complexes and model validation

Docking simulations of the YLOV dimer were carried out at the ClusPro Server,[47] using the DOT 1.0[48] and ZDOCK v. 2.3[49] programmes, employing the previously published YLOV structural model (PDB databank accession code 1IUM).[19] The ClusPro docking algorithm evaluates billions of putative complexes, retaining a preset number with favourable surface complementarities. A filtering method is then applied to this set of structures, selecting those with good electrostatic and desolvation free energies for further clustering. Evaluation of the ClusPro predicted complexes was carried out using the VADAR tool,[50] which calculates multi-chain parameters such as the accessible surface area and the percentage of accessible hydrophobic side chains. The quality of the ClusPro predicted YLOV dimer model was verified by using the protein structure analysis and validation server (SAVS) of the NIH MBI Laboratory for Structural Genomics and Proteomics at the University of California, Los Angeles (UCLA) (http://nihserver.mbi.ucla.edu/SAVS/) that implements the programs PROCHECK,[51] WHAT_CHECK,[52] ERRAT,[53] VERIFY_3D,[54] and PROVE.[55] Additionaly, PPI-Pred[56] and the computational interface alanine scanning tool at the Robetta server[57] were used to predict the dimerization interface of the YLOV-monomer and of the YLOV dimers respectively.

Results and discussion

Gel filtration chromatography

The elution profile of YLOV (Fig. 1) reveals that the majority of the protein is present in a state having $M_w = 36.44$ kDa, namely 2.63 times larger than the theoretical M_w (13.81 kDa, considering also the 20 aa at the N-terminal His-tag). This suggests a dimeric state that deviates from a spherical shape, as previously reported for other LOV domains.[30,32,33] Traces of a globular monomer are observable in some preparations, as well as a small fraction of a larger aggregate, most probably a trimer, with $M_w = 51.97$ kDa. Light activation does not appreciably affect the elution profile of YLOV. N-YLOV is also mostly present in a dimeric state, with apparent $M_w = 34.00$ kDa, 2.06 times larger than the theoretical M_w (16.54 kDa, again including the His-tag). For N-YLOV the shape of the dimer is therefore approximately spherical.

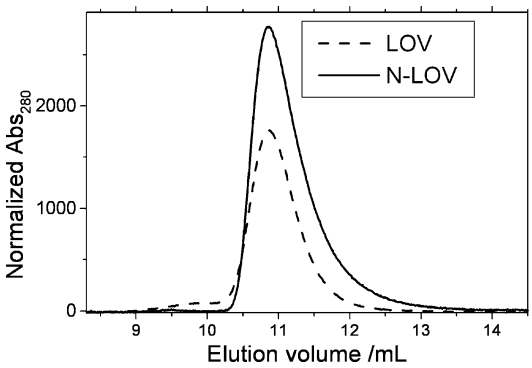


Fig. 1 Elution profiles of the YLOV (dashed line) and N-YLOV (full line) domains of YtvA. The main peaks correspond to $M_w = 2.6 \times M_{wYLOV}$ and $M_w = 2.1 \times M_{wN\text{-}YLOV}$, respectively.

The full-length protein YtvA presents a larger heterogeneity and the elution profile is different among different preparations (Fig. 2). Up to three peaks can be identified, with peak 1 resulting in $M_w = 72.85$ kDa, peak 2 with $M_w = 48.39$ kDa and peak 3 with $M_w = 35.07$ kDa. The theoretical M_w of YtvA is 31.36 kDa, therefore we can assign the three peaks to a dimeric state (peak 1), an elongated monomer (peak 2, $M_w = 1.56 \times M_{wYtvA}$) and a spherical monomer (peak 3, $M_w = 1.12 \times M_{wYtvA}$). Peak 2 (elongated monomer) represents in all YtvA preparations observed

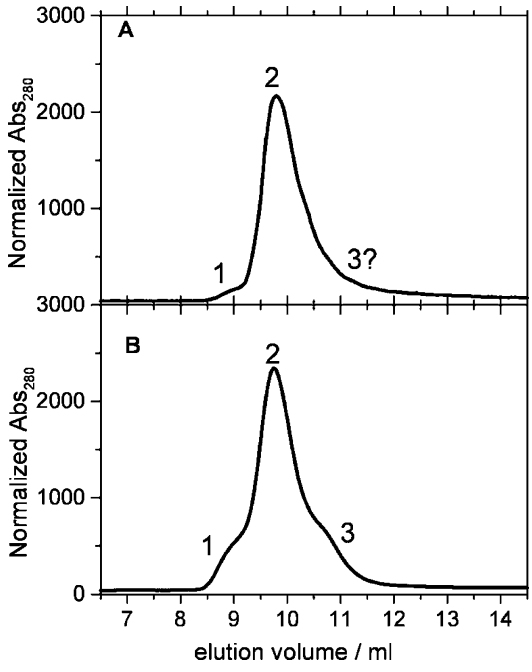


Fig. 2 Elution profile of 2 different YtvA preparations. Peak 1 corresponds to $M_w = 2.32 \times M_{wYtvA}$ ($M_{wYtvA} = 31.36$ kDa); peak 2 corresponds to $M_w = 1.56 \times M_{wYtvA}$; peak 3 corresponds to $M_w = 1.12 \times M_{wYtvA}$. A. YtvA : FMN = 1 : 1 (no apoprotein). B. YtvA : FMN ≈ 2 : 1 (apoprotein is present).

(nine in total) the predominant protein fraction. Peaks 1 and 3 are more evident in preparations that contain considerable amount of apoprotein (without the flavin chromophore). Nevertheless, even in these cases, the flavin chromophore is present in all three fractions, as proven by detection at 390 nm (data not shown). Again light activation does not result in appreciable changes in the elution profile. For the three protein constructs, the elution profile is not affected by concentration, in the low range employed here (0.05–2.5 μM).

As a whole, the gel filtration experiments show that the LOV domain of YtvA has a strong tendency to dimerize in solution that is not hindered by the N-terminal cap. Dimerization is instead prevented by the presence of the C-terminal domain of the protein, pointing to the fact that the LOV core employs the same surface (partially or totally) for homodimerization and for interdomain interactions.

Circular dichroism spectroscopy

The UV-CD spectra for the three analyzed constructs of YtvA, in the dark adapted state, are shown in Fig. 3. The mean residue ellipticity, Θ_{MRW}, was calculated as explained in the experimental section. The spectra shown are an average of all measurements (5 sets of measurements with 2 different preparations for YLOV; 4 sets of measurements with 2 different preparations for N-YLOV; 11 sets of measurements, 9 different preparations for YtvA), but each single spectral output was analyzed separately using the CCA algorithm.

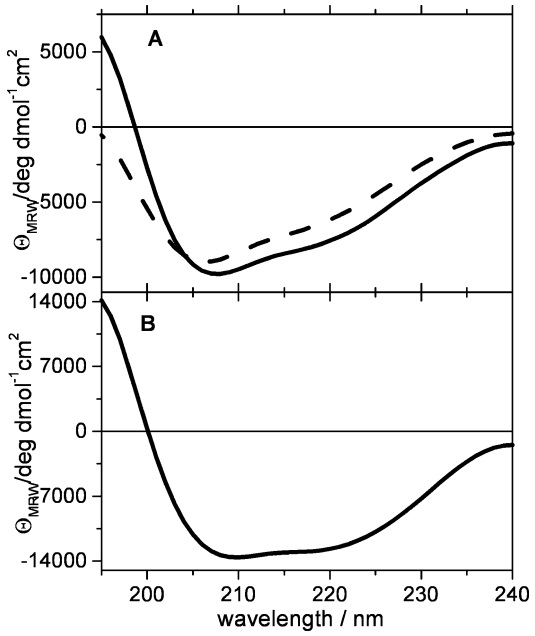


Fig. 3 CD spectra in the UV region for (A) the YLOV (dashed line) and N-YLOV (full line) domains and (B) full-length YtvA, in the dark adapted state.

A critical step during CCA analysis of CD data, is the assignment of the component curves (Fig. 4) to specific secondary structures. In the literature there is a large agreement about the CD spectrum of regular α-helices and unordered polypeptides

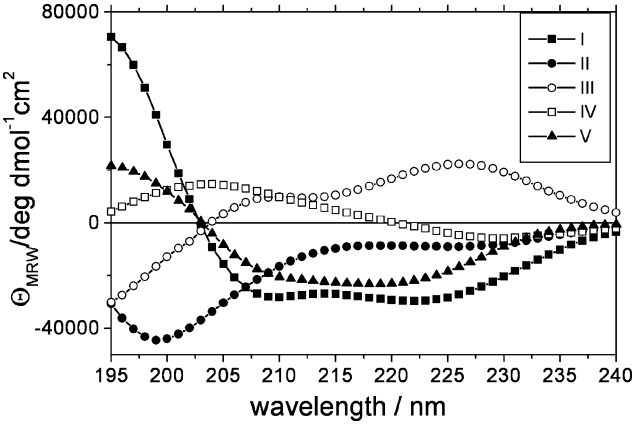


Fig. 4 The 5 component curves as extracted from the CCA analysis. The assignment is as follows: I, α-helix, II, unordered structures, III, turns and other structures, IV, distorted/twisted β-strands/parallel β-sheet, V, antiparallel β-sheet (see text for details).

(random coil, RC) that can be assigned to curve I and II respectively (ref. 58 and 59and references therein).

Curve V was assigned to the turn fraction[60] and the β-structures were assigned to curve IV + V, whose sum is similar to the curve corresponding to β-strands in Matsuo *et al.*[60] Based on the fact that distorted/twisted β-sheets present a strong positive band in the 190–220 region,[61] curve IV is assumed to include this fraction. Furthermore, the antiparallel β-sheet has three allowed transitions (the ππ* transition is splitted),[62,63] whereas the parallel β-sheet has two; therefore, we assigned curve V to the regular antiparallel β-sheet and curve IV to twisted + parallel β-sheet. Curve IV may also receive contributions from turn structures, in that different type of turns have very different CD spectra.[44,64,65] The results of the CCA analysis, with the curve assignment as discussed above, are reported in Table 1, both for the dark and light-adapted state.

To test the quality/reliability of our component assignment, we made a prediction of secondary structure composition based on the three dimensional models of the LOV and STAS domains[19,39] and on the consensus method for the remaining parts of the protein.[46] The N-cap and the linker region are predicted to be largely helical, whereas the His-tag (20 aa in length) is, as expected, predicted to be unordered (Table 2).

The comparison with CD data is very good in the case of the helical fraction, although for YtvA the statistical error associated with this component is quite large. This may be due to the variability in the preparations, and/or to the fact that component I and V are in some cases difficult to separate (see Fig. 4 and Table 1). The results confirm that the N-cap and the linker region are mostly helical. For full-length YtvA also the turn/loops and β-strands predicted fractions match the sum of component II + III and IV + V, respectively, within the experimental error, (Table 2), supporting our curve assignment. In the case of LOV and, particularly, N-LOV, the fraction of β-strands is smaller than expected, to the advantage of the RC/turns/others component. Furthermore, LOV domains do not contain parallel β-sheets[19,22,23] and the large percentage associated to component IV has to be assigned to the distortion/twisting of the central, antiparallel β-scaffold. This is in contrast with YtvA, for which the number of aa associated to component IV, can be readily explained with the presence of four parallel β-strands localized on the STAS

Table 1 Results of the CCA analysis on CD spectra

	LOV[a] (122 aa)[b] (%)		N-LOV[a] (147 aa)[b] (%)		YtvA[a] (281 aa)[b] (%)	
Secondary structure	Dark	Light	Dark	Light	Dark	Light
I, α-Helix	16.9 ± 2.7 (21 ± 3)	19.5 ± 5.8 (24 ± 7)	24.9 ± 2.4 (37 ± 4)	27.3 ± 0.9 (40 ± 1)	30.9 ± 4.7 (87 ± 13)	32.0 ± 6.5 (90 ± 18)
II, RC	23.9 ± 4.1 (29 ± 5)	22.8 ± 3.9 (28 ± 5)	24.9 ± 0.9 (37 ± 1)	23.6 ± 0.9 (35 ± 1)	22.4 ± 1.0 (63 ± 3)	21.7 ± 3.5 (61 ± 10)
III, β-Turns/others	27.9 ± 1.1 (34 ± 1)	28.8 ± 2.3 (35 ± 3)	27.5 ± 1.0 (40 ± 1)	28.3 ± 1.5 (42 ± 2)	16.8 ± 3.6 (47 ± 10)	18.7 ± 3.5 (52 ± 10)
IV, β-Twisted/β-parallel	14.7 ± 1.6 (18 ± 2)	18.6 ± 1.0 (23 ± 1)	12.1 ± 1.6 (18 ± 2)	15.9 ± 0.9 (23 ± 1)	9.9 ± 4.1 (28 ± 11)	11.0 ± 3.7 (31 ± 10)
V, β-Antiparallel	16.2 ± 4.1 (20 ± 5)	10.2 ± 4.7 (12 ± 6)	10.6 ± 3.4 (15 ± 5)	4.9 ± 2.1 (7 ± 3)	19.9 ± 5.6 (56 ± 16)	16.6 ± 5.8 (47 ± 16)
$\left\langle \sum_{i=1}^{n} [y_i - f(\lambda)]^2 \right\rangle^c$	1.9 ± 0.4	3.6 ± 2.9	2.2 ± 1.8	2.1 ± 0.7	6.2 ± 2.8	5.2 ± 2.3

[a] The statistical error is the standard deviation and comes from 5 sets of measurements on 2 different preparations for LOV, 4 sets of measurements on 2 different preparations for N-LOV, 11 sets of measurements, and 9 different preparations for YtvA. [b] The number of aa is given in parentheses, below the percentage, together with the statistical error. [c] Average squared error, where y_i = experimental curve, $f(\lambda)$ = fitting curve.

Table 2 Comparison between expected and CD-derived secondary structure composition

YtvA segments	Helices Number of aa	Turns/loops Number of aa	β-Strands *Numberof aa*
His-Tag[a]	—	20	—
N-Cap$_{1-24}$[a]	10	13	1
LOV$_{25-126}$[b]	24	36	42
Linker$_{127-146}$[a]	18	2	—
STAS$_{147-254}$[b]	36	41	31
C-End$_{255-261}$[a]	2	4	1
YtvA[c]	90	116	75
N-LOV[c]	34	69	43
LOV[c]	24	56	42
CCA analysis—dark state	**α-Helix**	**RC/turns/others**	**β-Strands**
YtvA	87 ± 13	110 ± 9	84 ± 17
N-LOV	37 ± 4	77 ± 1	33 ± 5
LOV	21 ± 3	63 ± 4	38 ± 5

[a] Consensus secondary structure prediction at the Pôle Bioinformatique Lyonnais server.[46] [b] Structural homology models of the LOV core[19] and of the STAS domain.[39] [c] Predicted number of aa for each of the constructs analyzed.

domain and a modest distortion of the overall β-fraction.[39] These observations suggest that dimerization in YLOV and N-YLOV markedly affects the central β-sheet of the LOV core (see the docking section).

Inspection of Table 1 and of the light − dark difference spectra (Fig. 5) shows that light activation of the three analyzed constructs does not result in large secondary structure conformational changes, as previously noticed for full-length YtvA.[41] A further distortion of the central β-sheet is induced in YLOV and N-YLOV, and the difference spectra are very similar for the two proteins.

Light-induced changes of the central β-sheet have been recently demonstrated with low temperature Fourier transformed infrared spectroscopy (FTIR) also for phy3-LOV2.[66] The CD light-difference spectrum of phot1-LOV2 was interpreted as a loss of helical structure, but without the support of a detailed data analysis.[67]

In full-length YtvA there is still a perturbation of the β-fraction, but a distinct change in the turn fraction (positive shoulder at *ca.* 230 nm in Fig. 5), missing in YLOV and N-YLOV. The determinations are affected by a large error, but confirmed by previous data as obtained with FTIR, that show a distinct

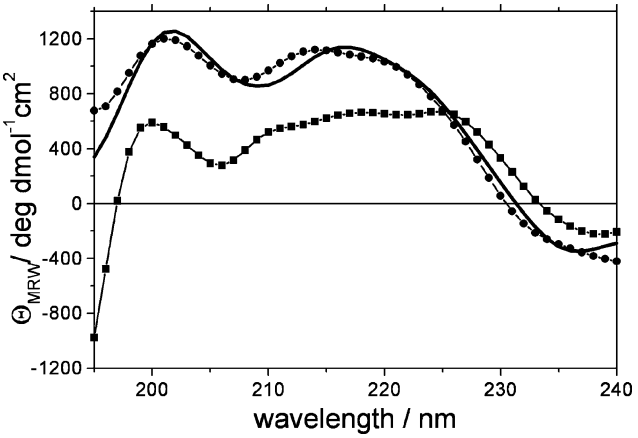


Fig. 5 Light − dark plots of the mean residue ellipticity Θ_{MRW} for YtvA (squares + line), LOV (circles + line) and N-LOV (full line), calculated from the average CD spectra (see Fig. 1 and Table 1).

difference between YtvA and YLOV in the light-induced changes of the turns fraction (around 1700 cm^{-1}).[68] These results could be

interpreted as a conformational change transmitted from the LOV core to the STAS domain, but actually we have no hint to localize precisely the position of the altered turn fraction, that could even be on the LOV domain itself and its changes being not detectable in the LOV dimers. Temperature-dependent FTIR experiments show indeed that changes in the turn fraction occur before the conformational alterations of the β-sheet in phy3-LOV2, the latter changes only detectable at room temperature.[66] In this view, the light-induced conformational changes could reach the β-scaffold only in the YLOV and N-YLOV dimers and be limited by the presence of the second domain, so that the changes on the turn fraction can persist during the lifetime of the adduct in YtvA.

The YLOV–YLOV dimer

The ClusPro best ranked model (ZDOCK generated) is shown in Fig. 6. A very similar model is ranked at the first position by using the DOT docking software (not shown).

The evaluation of the complexes performed with the VADAR tool,[50] reveals that this model has a quite large buried surface area as well as a high percentage of buried hydrophobic side-chains (33.76%) (see electronic supplementary information, ESI†). This feature would favour dimerization in an aqueous environment, and agrees with the fact that YLOV is a stable dimer in solution, even at very low concentrations. The quality of the model, evaluated at the SAVS server (http://nihserver.mbi.ucla.edu/SAVS/) was as an overall good, with 84.1% of residues in most favoured regions, 12.5% of residues in additionally allowed region, 2.3% of residues in generously allowed regions and only 1.1% of residues in disallowed regions of the Ramachandran plot. The Verify-3D[54] score (96.1) and the Errat-quality factor[53] (98.9) were very high, both indicative of a reasonable and good resolved model-structure. Finally, the dimerization surface predicted with PPI-Pred[56] and Robetta,[57] identified high scoring regions for YLOV–YLOV interactions within Aβ, Bβ, Hβ and Iβ strands and the Hβ–Iβ loop (Fig. 6).

In the dimer models of Fig. 6, the two YLOV domains face each other with the central β-sheet, presenting an antiparallel mirror symmetry. The interface is mostly stabilized by hydrophobic interactions. This feature is not in common with phy3-LOV2, where a bunch of charged/polar amino acids forms an extended HB (hydrogen bonds) network with the corresponding residues on the second monomer, centered around His1011, Gln1013 (Hβ) and Asp1017 (Hβ–Iβ loop). Interestingly, His1011 and Gln1013 of LOV2 domains (Thr and A/T respectively on LOV1 domains), and this feature may account for the fact that LOV1 has a stronger tendency to dimerize than LOV2 in an aqueous environment.[30] Although the dimerization of phy3-LOV2 and the specific orientation of the two monomers may be an artifact of crystallization, a complex very similar to the phy3-LOV dimer is readily obtained by the docking algorithm (not shown) and the residues at the interface are part of the hot spots predicted by PPI-Pred and Robetta (Fig. 6). We note that for YLOV, complexes with similar orientation as phy3-LOV are also detected by the ClusPro docking algorithm (cluster 2 and 8 in the DOT output and cluster 10 in the ZDOCK output, see ESI†). Our choice of the model in Fig. 6 (cluster 1 for both DOT and ZDOCK outputs) is based on the ClusPro ranking, on the high surface complementarity and interactions symmetry, and on the presence of a cluster of hydrophobic amino acids at the interface, that nicely accounts for the stability of the dimer in solution.

The antiparallel mirror symmetry and the interface observed in our YLOV–YLOV model and in the phy3-LOV dimer is very similar to the one reported for homo and heterodimers of the ARNT PAS-B domain in solution[69] and in dimers of the heme-binding PAS domain of *E. coli* Dos (EcDos)[70] and *R. meliloti* FIXL (RmFIXL).[71] An antiparallel mirror symmetry has also been suggested for the LOV–LOV dimer of the FKF1 protein and for phot-LOV1 on the basis of small-angle X-ray scattering experiments, although in that case the authors favoured a different model for the complex, where the two LOV domains do not face each other *via* the central β-sheet.[32,33]

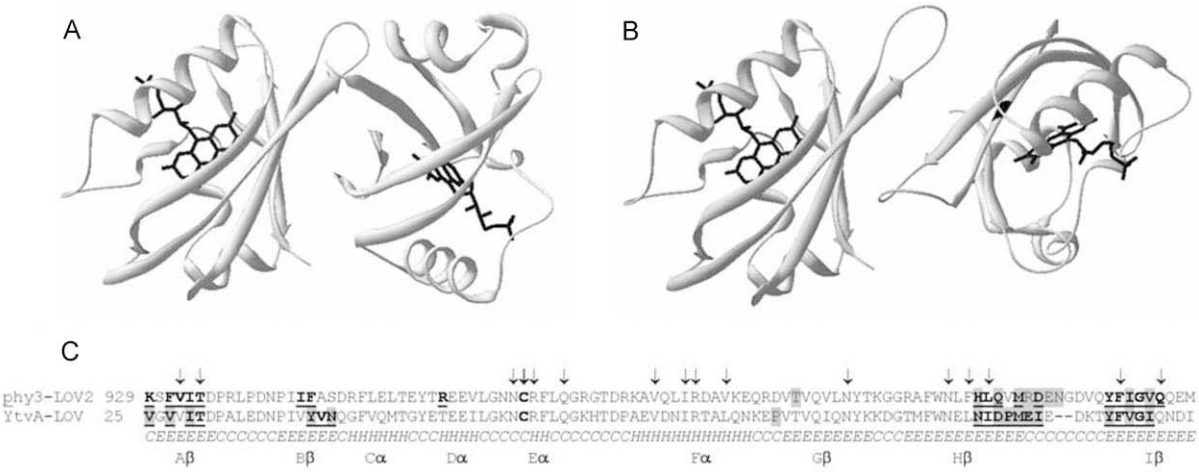


Fig. 6 (A) The YLOV dimer model (see text for details). (B) The phy3-LOV2 dimer in the crystal unit cell (PDB accession code 1G28, chains a,c).[22] (C) Residues at the dimer interface (within 4 Å, shadowed), mapped on the sequence of YLOV and phy3-LOV2. For comparison the interaction hot-spots predicted by PPI-Pred[56] (in bold) and by Robetta[57] are also shown (in bold and underlined). Arrows indicate the residues interacting with FMN in phy3-LOV2. Secondary structure elements are shown below the phy3-LOV2/YLOV alignment and indicated with conventional letters. E = strands, H = helices, C = unordered. The nomenclature of the secondary structure elements is after Harper *et al.*[24]

The model as in Fig. 6 not only corresponds to the best quality/validation parameters, but also agrees with the observations that in YLOV the β-scaffold is distorted/twisted within the dimer, as indicated by the CD data. With this respect we have to remind that the docking simulation requires that the partners within the complex are kept rigid, a feature that may not be verified in the real complex, as again suggested by the distortion of the β-scaffold. Therefore, the model structure reported in Fig. 6 has to be taken with care and, albeit probably qualitatively correct, may not match the solution dimer-structure as far as the details are concerned.

In the structure of EcDos and RmFIXL PAS domains, the dimers are further stabilized by the helical N-cap and the dimers retain an elongated shape.[70,71] In the case of N-YLOV, although the N-cap is helical (from CD data), the dimer is instead approximately spherical (gel filtration). This observation, together with low similarity to the corresponding sequences in EcDos and RmFIXL, does not allow to build a reliable model of the N-cap in YLOV and of its orientation with respect to LOV core.

Similar considerations apply to a structural model of the full-length protein. The LOV-STAS linker region is predicted to be helical and CD data confirm the prediction (see above), nevertheless we cannot safely state that it assumes an orientation similar to that in phot1-LOV2, namely underneath the central β-scaffold of the LOV core,[24] because of low sequence similarity. The amino acid sequence of YtvA Jα-linker is much more similar to the C-terminal extension of the heme binding PAS domain of FIXL from *B. japonicum*, actually protruding outside the PAS core.[72] Although this might be, in the latter case, an artifact of crystallization (in the absence of the associated kinase domain), such orientation of the Jα-linker cannot be excluded. This would imply a direct interaction of the STAS domain with the LOV core, competing with the dimerization surface, different to the structural features proposed for phot as a basis for the self-phosphorylation reaction.[24] We must also consider that alternative complex conformations may exist: the linker is not needed for the activation of the kinase activity in phot2 towards a substrate, a reaction carried out *via* direct interaction between the separately expressed LOV2 and kinase domains.[15] As a whole we still have too little structural and functional information about the N-cap and Jα-linker to build a reliable model of full-length YtvA. In order to gain further structural information, *e.g.*, orientation of the helical linker with respect to the LOV core and its relevance in the LOV-STAS interaction, we are designing separated constructs for the STAS domain and the LOV core furnished with the linker region.

We wish to point out that with CD experiments we can only see modifications in the secondary structure elements, but protein movements could occur without large conformational changes of the secondary structure. Furthermore our experiments are not time resolved, therefore we cannot detect transient structural changes occurring within the time-scale for the formation of the adduct (*ca.* 2 μs).[19]Another problem is represented by the fact that we are working with a system that only partially resembles physiological conditions. In fact, from a very recent paper we know that YtvA functions within a large macromolecular complex, about which we presently do not have any structural information.[37] Some hints about the way the STAS domain is activated may come from our recent experiments showing that YtvA binds Nucleotide TriPhosphate (NTP = GTP, ATP)[39] and that light-induced conformational changes are transmitted from the LOV-core to the NTP binding cavity on the STAS domain.[39]These changes are very small and certainly do not imply large structural changes in YtvA, but may have a larger significance within the macromolecular complex mentioned above.

Conclusions

In this work we have investigated the conformation of YtvA in solution. The analysis of CD spectra by means of the CCA algorithm and curve assignment has been improved with respect to previous work and can be now reliably employed to determine the secondary structural composition of LOV proteins. The LOV domain of YtvA has been proven to be an elongated dimer, stabilized by interactions that involve the β-scaffold, for which we have modelled a structure that agrees with experimental data and bioinformatic analysis. In the N-YLOV construct, the helical N-terminal cap is expected to participate in dimerization, although we could not model the complex due to lack of information about the orientation of this segment. In the full-length protein YtvA, dimerization appears only in case that apoprotein is present, most probably with the formation of heterodimers (apoprotein/FMN-bound YtvA). The data strongly suggest that the β-scaffold is involved both in YLOV dimerization and intraprotein interactions with the linker and/or STAS domain, confirming that this region is a good candidate as a surface responsible for signal transmission to the effector domains in LOV proteins. This latter aspect highlights a sharp similarity with phot-LOV2,[24] but the missing light-driven unfolding of the Jα-linker also points to a distinct difference between phot and YtvA, in the way the effector domain is activated. Furthermore, the data presented here suggest that dimerization of LOV domains might play an important regulative role by competing with domain–domain interactions and should be thoroughly investigated.

References

1 M. H. Hefti, K. J. Francoijs, S. C. de Vries, R. Dixon and J. Vervoort, The PAS fold: A redefinition of the PAS domain based upon structural prediction, *FEBS J.*, 2004, **271**, 1198–1208.

2 W. R. Briggs, C. F. Beck, A. R. Cashmore, J. M. Christie, J. Hughes, J. A. Jarillo, T. Kagawa, H. Kanegae, E. Liscum, A. Nagatani, K. Okada, M. Salomon, W. Rüdiger, T. Sakai, M. Takano, M. Wada and J. C. Watson, The phototropin family of photoreceptors, *Plant Cell*, 2001, **13**, 993–997.

3 W. R. Briggs and J. M. Christie, Phototropins 1 and 2: versatile plant blue-light receptors, *Trends Plant Sci.*, 2002, **7**, 204–210.

4 R. Banerjee and A. Batschauer, Plant blue-light receptors, *Planta*, 2005, **220**, 498–502.

5 E. Huala, P. W. Oeller, E. Liscum, I. S. Han, E. Larsen and W. R. Briggs, *Arabidopsis* NPH1: A Protein Kinase with a Putative Redox-Sensing Domain, *Science*, 1997, **278**, 2120–2123.

6 E. Knieb, M. Salomon and W. Rüdiger, Autophosphorylation, electrophoretic mobility and immunoreaction of oat Phototropin 1 Under UV and Blue Light, *Photochem. Photobiol.*, 2005, **81**, 177–182.

7 M. Salomon, J. M. Christie, E. Knieb, U. Lempert and W. R. Briggs, Photochemical and mutational analysis of the FMN-binding domains of the plant blue light receptor phototropin, *Biochemistry*, 2000, **39**, 9401–9410.

8 M. Salomon, W. Eisenreich, H. Dürr, E. Scleicher, E. Knieb, V. Massey, W. Rüdiger, F. Müller, A. Bacher and G. Richter, An optomechanical transducer in the blue light receptor phototropin from *Avena sativa*, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 12357–12361.

9 M. Kasahara, T. E. Swartz, M. A. Olney, A. Onodera, N. Mochizuki, H. Fukuzawa, E. Asamizu, S. Tabata, H. Kanegae, M. Takano, J. M.

Christie, A. Nagatani and W. R. Briggs, Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from *Arabidopsis*, rice, and *Chlamydomonas reinhardtii*, *Plant Physiol.*, 2002, **129**, 762–773.
10 S. Crosson and K. Moffat, Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch, *Plant Cell*, 2002, **14**, 1067–1075.
11 C. W. M. Kay, E. Schleicher, A. Kuppig, H. Hofner, W. Rüdiger, M. Schleicher, M. Fischer, A. Bacher, S. Weber and G. Richter, Blue light perception in plants. Detection and characterization of a light-induced neutral flavin radical in a C450A mutant of phototropin, *J. Biol. Chem.*, 2003, **278**, 10973–10982.
12 J. T. M. Kennis, S. Crosson, M. Gauden, I. H. M. van Stokkum, K. Moffat and R. van Grondelle, Primary Reactions of the LOV2 Domain of Phototropin, a Plant Blue-Light Photoreceptor, *Biochemistry*, 2003, **42**, 3385–3392.
13 T. E. Swartz, S. B. Corchnoy, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs and R. A. Bogomolni, The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin, *J. Biol. Chem.*, 2001, **276**, 36493–36500.
14 T. Kottke, J. Heberle Dominic Hehn and P. Hegemann, Phot-LOV1: Photocycle of a Blue-Light Receptor Domain from the Green Alga *Chlamydomonas reinhardtii*, *Biophys. J.*, 2003, **84**, 1192–1201.
15 D. Matsuoka and S. Tokutomi, Blue light-regulatedmolecular switch of Ser/Thr kinase in phototropin, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 13337–13342.
16 J. M. Christie, T. E. Swartz, R. A. Bogomolni and W. R. Briggs, Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function, *Plant J.*, 2002, **32**, 205–219.
17 S. Crosson, S. Rajagopal and K. Moffat, The LOV domain family: photoresponsive signaling modules coupled to diverse output domains, *Biochemistry*, 2003, **42**, 2–10.
18 A. Losi, The bacterial counterparts of plants phototropins, *Photochem. Photobiol. Sci.*, 2004, **3**, 566–574.
19 A. Losi, E. Polverini, B. Quest and W. Gärtner, First evidence for phototropin-related blue-light receptors in prokaryotes, *Biophys. J.*, 2002, **82**, 2627–2634.
20 U. Krauss, A. Losi, W. Gärtner, K.-E. Jaeger and T. Eggert, Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct, *Phys. Chem. Chem. Phys.*, 2005, **7**, 2229–2236.
21 E. Schleicher, R. M. Kowalczyk, C. W. M. Kay, P. Hegemann, A. Bacher, M. Fischer, R. Bittl, G. Richter and S. Weber, On the reaction mechanism of adduct formation in LOV domains of the plant blue-light receptor phototropin, *J. Am. Chem. Soc.*, 2004, **126**, 11067–11076.
22 S. Crosson and K. Moffat, Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 2995–3000.
23 R. Fedorov, I. Schlichting, E. Hartmann, T. Domratcheva, M. Fuhrmann and P. Hegemann, Crystal structures and molecular mechanism of a light-induced signaling switch: the Phot-LOV1 domain from *Chlamydomonas reinhardtii*, *Biophys. J.*, 2003, **84**, 2492–2501.
24 S. M. Harper, L. C. Neil and K. H. Gardner, Structural basis of a phototropin light switch, *Science*, 2003, **301**, 1541–1544.
25 S. M. Harper, J. M. Christie and K. H. Gardner, Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity, *Biochemistry*, 2004, **43**, 16184–16192.
26 A. Losi, in *Flavin photochemistry and photobiology*, ed. D.-P. Häder and G. Jori, Elsevier, Amsterdam, 4th edn, 2006, ch. 10, pp. 223–276.
27 T. Eitoku, Y. Nakasone, D. Matsuoka, S. Tokutomi and M. Terazima, Conformational dynamics of phototropin 2 LOV2 domain with the linker upon photoexcitation, *J. Am. Chem. Soc.*, 2005, **127**, 13238–13244.
28 B. L. Taylor and I. B. Zhulin, PAS domains: internal sensors of oxygen, redox potential and light, *Microbiol. Mol. Biol. Rev.*, 1999, **63**, 479–506.
29 M. A. Gilles-Gonzalez and G. Gonzalez, Heme-based sensors: defining characteristics, recent developments, and regulatory hypotheses, *J. Inorg. Biochem.*, 2005, **99**, 1–22.
30 M. Salomon, U. Lempert and W. Rüdiger, Dimerization of the plant photoreceptor phototropin is probably mediated by the LOV1 domain, *FEBS Lett.*, 2004, **572**, 8–10.
31 Y. Nakasone, T. Eitoku, D. Matsuoka, S. Tokutomi and M. Terazima, Kinetic Measurement of Transient Dimerization and Dissociation Reactions of Arabidopsis Phototropin 1 LOV2 Domain, *Biophys. J.*, 2006, **91**, 645–653.
32 M. Nakasako, D. Matsuoka, K. Zikihara and S. Tokutomi, Quaternary structure of LOV-domain containing polypeptide of *Arabidopsis* FKF1 protein, *FEBS Lett.*, 2005, **579**, 1067–1071.
33 M. Nakasako, T. Iwata, D. Matsuoka and S. Tokutomi, Light-Induced Structural Changes of LOV Domain-Containing Polypeptides from *Arabidopsis* Phototropin 1 and 2 Studied by Small-Angle X-ray Scattering, *Biochemistry*, 2004, **43**, 14881–1489.
34 P. Ballario, C. Talora, D. Galli, H. Linden and G. Macino, Roles in dimerization and blue light photoresponse of the PAS and LOV domains of *Neurospora crassa* white collar proteins, *Mol. Microbiol.*, 1998, **29**, 719–729.
35 L. Aravind and E. V. Koonin, The STAS domain a link between anion transporters and antisigma-factor antagonists, *Curr. Biol.*, 2000, **10**, R53–R55.
36 S. Akbar, T. A. Gaidenko, K. Min, M. O'Reilly, K. M. Devine and C. W. Price, New family of regulators in the environmental signaling pathway which activates the general stress transcription factor of *Bacillus subtilis*, *J. Bacteriol.*, 2001, **183**, 1329–1338.
37 T. A. Gaidenko, T. J. Kim, A. L. Weigel, M. S. Brody and C. W. Price, The blue-light receptor YtvA acts in the environmental stress signaling pathway of *Bacillus subtilis*, *J. Bacteriol.*, 2006, **188**, 6387–6395.
38 M. Avila-Perez, K. J. Hellingwerf and R. Kort, Blue light activates the sigmaB-dependent stress response of *Bacillus subtilis via* YtvA, *J. Bacteriol.*, 2006, **188**, 6411–6414.
39 V. Buttani, A. Losi, E. Polverini and W. Gärtner, Blue news: NTP binding properties of the blue-light sensitive YtvA protein from *Bacillus subtilis*, *FEBS Lett.*, 2006, **580**, 3818–3822.
40 S. M. Najafi, D. A. Harris and M. D. Yudkin, The SpoIIAA protein of *Bacillus subtilis* has GTP-binding properties, *J. Bacteriol.*, 1996, **178**, 6632–6634.
41 A. Losi, E. Ghiraldelli, S. Jansen and W. Gärtner, Mutational effects on protein structural changes and interdomain interactions in the blue-light sensing LOV protein YtvA, *Photochem. Photobiol.*, 2005, **81**, 1145–1152.
42 A. Losi, B. Quest and W. Gärtner, Listening to the blue: the time-resolved thermodynamics of the bacterial blue-light receptor YtvA and its isolated LOV domain, *Photochem. Photobiol. Sci.*, 2003, **2**, 759–766.
43 A. Perczel, M. Hollosi, G. Tusnady and G. D. Fasman, Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins, *Protein Eng.*, 1991, **4**, 669–679.
44 A. Perczel, K. Park and G. D. Fasman, Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: A practical guide, *Anal. Biochem.*, 1992, **203**, 83–93.
45 I. R. Bates, P. Matharu, N. Ishiyama, D. Rochon, D. D. Wood, E. Polverini, M. A. Moscarello, N. J. Viner and G. Harauz, Characterization of a Recombinant Murine 18.5-kDa Myelin Basic Protein, *Protein Expression Purif.*, 2000, **20**, 285–299.
46 C. Combet, C. Blanchet, C. Geourjon and G. Deleage, NPS@: Network Protein Sequence Analysis, *Trends Biochem. Sci.*, 2000, **25**, 147–150.
47 S. R. Comeau, D. W. Gatchell, S. Vajda and C. J. Camacho, ClusPro: An automated docking and discrimination method for the prediction of protein complexes, *Bioinformatics*, 2004, **20**, 45–50.
48 J. G. Mandell, V. A. Roberts, M. E. Pique, V. Kotlovyi, J. C. Mitchell, E. Nelson, I. Tsigelny and L. F. Ten Eyck, Protein docking using continuum electrostatics and geometric fit, *Protein Eng.*, 2001, **14**, 105–113.
49 R. Chen, L. Li and Z. Weng, ZDOCK: an initial-stage protein-docking algorithm, *Proteins*, 2003, **52**, 80–87.
50 L. Willard, A. Ranjan, H. Zhang, H. Monzavi, R. F. Boyko, B. D. Sykes and D. S. Wishart, VADAR: a web server for quantitative evaluation of protein structure quality, *Nucleic Acids Res.*, 2003, **31**, 3316–3319.
51 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, PROCHECK-A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
52 R. W. Hooft, G. Vriend, C. Sander and E. E. Abola, Errors in protein structures, *Nature*, 1996, **381**, 272–272.
53 C. Colovos and T. O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, *Protein Sci.*, 1993, **2**, 1511–1519.
54 R. Luthy, J. U. Bowie and D. Eisenberg, Assessment of protein models with three-dimensional profiles, *Nature*, 1992, **356**, 83–85.
55 J. Pontius, J. Richelle and S. J. Wodak, Deviations from standard atomic volumes as a quality measure for protein crystal structures, *J. Mol. Biol.*, 1996, **264**, 121–136.

56 J. R. Bradford and D. R. Westhead, Improved prediction of protein-protein binding sites using a support vector machines approach, *Bioinformatics*, 2005, **21**, 1487–1494.
57 T. Kortemme, D. E. Kim and D. Baker, Computational alanine scanning of protein–protein interfaces, *Sci. STKE*, 2004, **2004**(219), pl2.
58 N. Sreerama, S. Y. Venyaminov and R. W. Woody, Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis, *Anal. Biochem.*, 2000, **287**, 243–251.
59 N. Sreerama, S. Y. Venyaminov and R. W. Woody, Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy, *Protein Sci.*, 1999, **8**, 370–380.
60 K. Matsuo, R. Yonehara and K. Gekko, Improved Estimation of the Secondary Structures of Proteins by Vacuum-Ultraviolet Circular Dichroism Spectroscopy, *J. Biochem.*, 2005, **138**, 79–88.
61 N. Sreerama and R. W. Woody, Structural composition of betaI- and betaII-proteins, *Protein Sci.*, 2003, **12**, 384–388.
62 N. J. Greenfield, Analysis of circular dichroism data, *Methods Enzymol.*, 2004, **383**, 282–317.
63 R. W. Woody and A. Koslowski, Recent developments in the electronic spectroscopy of amides and alpha-helical polypeptides, *Biophys. Chem.*, 2002, **101–102**, 535–551.
64 N. J. Greenfield, Methods to estimate the conformation of proteins and polypeptides from circular dichroism data, *Anal. Biochem.*, 1996, **235**, 1–10.
65 C. T. Chang, C. S. Wu and J. T. Yang, Circular dichroic analysis of protein conformation: inclusion of the beta-turns, *Anal. Biochem.*, 1978, **91**, 13–31.
66 T. Iwata, D. Nozaki, S. Tokutomi and H. Kandori, Comparative investigation of the LOV1 and LOV2 domains in *Adiantum* Phytochrome3, *Biochemistry*, 2005, **44**, 7427–7434.
67 S. B. Corchnoy, T. E. Swartz, J. W. Lewis, I. Szundi, W. R. Briggs and R. A. Bogomolni, Intramolecular proton transfers and structural changes during the photocycle of the LOV2 domain of phototropin 1, *J. Biol. Chem.*, 2003, **278**, 724–731.
68 T. Bednarz, A. Losi, W. Gärtner, P. Hegemann and J. Heberle, Functional variations among LOV domains as revealed by FT-IR difference spectroscopy, *Photochem. Photobiol. Sci.*, 2004, **3**, 575–579.
69 P. B. Card, P. J. A. Erbel and K. H. Gardner, Structural Basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization, *J. Mol. Biol.*, 2005, **353**, 664–677.
70 H. J. Park, C. Suquet, J. D. Satterlee and C. Kang, Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of *Escherichia Coli* Dos Heme Domain (*Ec*DosH), *Biochemistry*, 2004, **43**, 2738–2746.
71 H. Miyatake, M. Mukai, S. Y. Park, S. Adachi, K. Tamura, H. Nakamura, K. Nakamura, T. Tsuchiya, T. Iizuka and Y. Shiro, Sensory mechanism of oxygen sensor FixL from Rhizobium meliloti: crystallographic, mutagenesis and resonance Ramanspectroscopic studies, *J. Mol. Biol.*, 2000, **301**, 415–431.
72 W. Gong, B. Hao, S. S. Mansy, G. Gonzalez, M. A. Gilles-Gonzalez and M. K. Chan, Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 15177–15182.

Blue-light photoreceptors of the LOV-family in *Pseudomonas putida:* a conserved family of sensor modules in saprotrophic Pseudomonads†

Ulrich Krauss,*[a] Aba Losi,[b,c] Thorsten Eggert.[d] Wolgang Gärtner[e] and Karl-Erich Jaeger[a]

Manuscript in Preparation

Pseudomonas putida **KT2440, a typical saprotrophic bacterium, bears two genes encoding for LOV-proteins, PpSB1-LOV and PbSB2-LOV, both harbouring a LOV-core with N- and C-terminal flanking regions and lacking a fused effector domain. For PpSB2-LOV we have previously demonstrated the occurrence of a typical LOV-photocycle. Here we present the spectroscopic characterization of PpSB1-LOV and show that the two proteins, although they are very similar in sequence, exhibit dramatically different dark recovery kinetics. This study was addressed to identify potential signaling pathways of the two LOV-domain proteins, and, more generally, the effects of blue light irradiation on the physiological behaviour of** ***P. putida*****. Bioinformatic analysis of the genome proximity for PpSB1-LOV and PpSB2-LOV, revealed a clustering of genes encoding for proteins involved in the iron-starvation response, prompting us to investigate a possible blue-light regulation of this process. We detected under blue-light and iron starvation conditions an approximately 2-fold increase in the production of the iron-siderophore pyoverdine with respect to cells kept in the dark or under green or red light. Furthermore, growth of** ***P. putida*** **cells under blue light irradiation or in darkness (control) gave first evidence that blue light might have a stimulating effect on biofilm formation. The blue-light sensitivity and the** ***in silico*** **analysis suggest that the two LOV proteins are involved in these processes. Due to the conservation of similar LOV photosensor modules in several saprotrophic Pseudomonads, we anticipate functionally conserved blue-light responses and signaling processes among this group of organisms.**

Introduction

One of the most ancient environmental stimuli and/or energy source that all organisms, even before the advent of oxygen in the primordial atmosphere some 2.3 billion years ago[1], encountered or relied on, was the light emitted by the sun. On the other hand, in particular the short wavelength region of the overall incoming radiation, i.e., the UV- and blue-light range can cause deleterious effects in all living beings by: i) directly causing cellular damage on proteins and DNA due to direct photon absorption and excited state formation in the affected molecular structures and cellular compartments (UV-light)[2] and ii) by the excitation of endo- and exogenous photosensitizing molecules such as porphyrins and flavins, to their corresponding triplet-states[3, 4]. Those excited photosensitizers can subsequently induce the formation of free radicals (by electron transfer and proton abstraction, Type I photosensitization) and/or energy transfer to a secondary acceptor (e.g. O_2)(Type II). In aerobic environments both Type I and Type II reactions effectively generate reactive oxygen species (ROS) among which singlet oxygen is one of the most powerful known oxidant agents[5].

[a] *Institut für Molekulare Enzymtechnologie, Heinrich-Heine Universität Düsseldorf, Forschungszentrum Jülich, D-52426 Jülich, Germany; E-mail: u.krauss@fz-juelich.de*
[b] *Department of Physics, University of Parma, via G.P. Usberti 7/A,, 43100-Parma, Italy*
[c] *CNR-INFM, Parma, Italy*
[d] *EVOcatal GmbH, Merowingerplatz 1a, D-40225 Düsseldorf, Germany*
[e] *Max-Planck-Institute for Bioinorganic Chemistry, Stiftstraße 34-36, D-45470 Mülheim an der Ruhr, Germany*

ROS are known to cause damage in a wide range of cellular tissues and organs[3] It is thus not surprising that many living beings have developed mechanisms to i) either utilize the incoming radiation efficiently by employing phototropism as in case of the plants[6] or positive phototaxis as in case of microbes[7], to optimize the photosynthetic / phototrophic processes, ii) employ mechanisms to repair the caused damage (e.g., by the action of DNA photolyases)[8] or iii) to avoid high light exposure by negative phototaxis or photophobic responses[7]. All these three processes are dependent on the ability of the organisms to sense the source of radiation and optimally integrate physical parameters such as wavelength, intensity and duration to produce a balanced physiological response. For this purpose many organisms are equipped with a variety of photoreceptors that rely on the binding of a light sensitive chromophore to "detect" a specific wavelength range of the visible light. Currently six different photoreceptor families that are widespread in pro-and eukaryotes are distinguished: The red/far-red sensing phytochromes[9, 10], sensory rhodopsins[11, 12], xanthopsins[13] as well as the blue-light sensing flavin-binding photoreceptors of the cryptochrome[14]-, the BLUF- (blue-light sensing using FAD)[15] and LOV- (light, oxygen, voltage) families[16-18]. One of the most widely distributed class of flavin-binding photoreceptors seems to be the LOV family[17].
The best characterized LOV domain-containing photoreceptors are the plant phototropins that contain invariably two LOV domains (LOV1 and LOV2), coupled to a serine threonine kinase, which is autophosphorylated upon blue-light exposure[6]. Phototropins (phot1 and phot2) were initially identified as the primary photoreceptors that mediate blue-light dependent plant phototropism[19]. However, it became

soon apparent that they control also other blue-light dependent phenomena such as chloroplast movement, leaf expansion, and stomata opening[20]. Likewise, in the green algae *Chlamydomonas reinhardtii*, phot is instead involved in the control of multiple steps in the sexual life cycle[21].

Due to an increasing number of biophysical and structural studies, the reaction mechanisms, by which LOV domains sense and react to blue-light are well understood. Furthermore, it became clear that the LOV photochemistry is conserved between eu- and prokaryotic taxa: photon absorption by the FMN molecule that is bound non-covalently in LOV domain dark state (LOV447), excites the FMN molecule on a ps- time scale[22] to its excited triplet-state ^{3}[FMN](LOV660). The signaling state (LOV390) formation occurs via the decay of the LOV660, typically with a lifetime of 1-2µs[23, 24]. LOV390 contains a FMN-cysteine (C4a) -thiol adduct[25-27], and the only amino acid residue indispensable for the adduct formation is a cysteine residue located in a sequence motif (GXN**C**RFLQG) conserved among the majority of LOV domains. The signaling state LOV390 is the longest living species of the LOV photocycle and thermally reconverts in the dark to the ground state within minutes to hours for plant and bacterial LOV domains[28-32]. This recovery can take up to several days, as reported for the LOV domain of the Flavin-Binding-Kelch-Repeat F-box protein (FKF1) of *Arabidopsis* [31].

The first prokaryotic LOV domain-containing protein, biochemically and biophysically characterized was the *Bacillus subtilis* YtvA protein[30]. Subsequently, LOV domain homologous sequences were identified in a variety of photosynthetic and non-photosynthetic prokaryotes. Consequently, photochemical studies on various prokaryotic LOV-proteins were conducted. Data is now available for LOV-proteins from *B. subtilis*[30], *P. putida*[28], *Caulobacter crescentus*[29], a (filamentous) cyanobacterium *Anabaena* sp. PCC 7120[33] and very recently also for a hybrid LOV-histidine kinase (LOV-HK-RR, whereas, RR stands for response regulator that C-terminally to the kinase) from *Pseudomonas syringae pv.* tomato[34, 35], a LOV-HK from the marine phototroph *Erythrobacter litoralis* and LOV-HKs from the intracellular animal/human pathogens *Brucella abortus* and *Brucella melitensis*[35]. However, their physiological importance, especially in non-photosynthetic soil microorganisms such as *B. subtilis* and *P. putida* is in most cases poorly understood.

Only recently, evidence was provided that YtvA mediates blue-light dependent physiological responses in *B. subtilis*. Gaidenko and co-workers were able to demonstrate that YtvA is part of large environmental-signaling complex that controls or modulates the activation of the *B. subtilis* general stress transcription factor σ^B [36] and that the cysteine62, substrate in the YtvA photocycle, is required for the biological response *in vivo*[36, 37]. Furthermore, it was shown by Avila-Perez and colleagues that the activation of the *B. subtilis* general stress response (governed by σ^B) is directly modulated by blue-light due to the action of YtvA[38]. Very recently, Swartz and co-workers were able to demonstrate that for the intracellular animal pathogen *B. abortus,* efficient cell-replication in its host and hence virulence is strongly stimulated by blue-light[35], and moreover, that the organisms LOV-HK is mediating this response.

Therefore, a few years after the identification of LOV domains in the plant phototropins, experimental evidence is brought forward that proves the functional conservation of the LOV paradigm between the pro- and eukaryotic domains of life. Nevertheless, the available information is still scarce compared to some plant LOV signaling systems (e.g., the circadian photoreceptors of the ZEITLUPE-LOV family[39]), for which detailed physiological responses but also signaling/regulatory cascades are described[40, 41].

We reported recently the photochemical and biochemical characterization of a LOV blue-light signaling module (termed PpSB2-LOV)[28], identified in the plant-root colonizing proteobacterium *P. putida* KT2440[42]. Here we present evidence that the same microorganism possesses a second LOV-protein (termed PpSB1-LOV), with a very similar sequence to PpSB2-LOV, but with a significantly different photochemical behavior with respect to dark state recovery kinetics. Based on bioinformatic analysis and initial physiological studies using *P. putida* KT2440, we demonstrate the influence of blue-light in the iron-starvation / stationary phase response in this organism, which implies the action of (a) blue-light receptor(s) and hence suggest a possible role for the two LOV photoreceptor proteins. Furthermore, the conservation of the LOV-protein containing gene-clusters among several saprotrophic (here root-colonizing) Pseudomonads suggests a conserved functional role for blue-light signaling in those organisms.

Experimental

Bioinformatic analysis.

Annotated sequences of the genome regions of *Pseudomonas* strains, for which BLAST analyses had suggested the presence of LOV domain containing sensory box proteins (either PpSB1-LOV and/or PpSB2-LOV-like), were obtained from the Genebank database at the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov).

Those *Pseudomonas* genome regions, were manually inspected for the conservation of operon-like structures or a clustering of certain functional proteins in the vicinity of the respective LOV-protein.

Amino acid sequence alignments were performed using the AlignX tool of the Vector-NTI software package (Invitrogen, Carlsbad, CA, USA). Alignments were visualized and edited using the alignment editor Genedoc[43] and pairwise sequence identities were derived using the same tool. The Genome proximity for the two sensory box proteins (PpSB1-LOV and PpSB2-LOV) of *P. putida* KT2440 was compared using the PutidaList server of the Institute Pasteur (http://bioinfo.hku.hk/GenoList/index.pl?database=putidalist). Phylogenetic analyses were performed using either PhyML[44], MultiPhyl[45] or IQPNNI[46]. Bootstrapped Maximum-Likelihood trees with 100 replicates for each

run were generated using the PhyML-Server (http://atgc.lirmm.fr/phyml/) and the MultiPhyl-server (http://distributed.cs.nuim.ie/multiphyl.php), respectively.
Next-Neighbour Interchange and Subtree Pruning and Recrafting heuristic tree-searches with 100 bootstrap replicates were performed using the MultiPhyl-server. Additionally, Bayesian posterior probability support values were added to each branch of the ML-tree using BEAST v1.4.5[47]. An initial run of ProtTest[48] on the input alignment, suggested a JTT+G model of amino acid substitution (with 4 gamma-shapes and an alpha value of 1.751) as best suiTable for reconstructing the sequence evolution in our dataset. Therefore, all phylogenetic tree reconstructions were carried out using this amino acid substitution matrix.
For Bayesian computations random starting trees were used, and analyses were run for 5 million generations, sampling the Markov chain Monte Carlo (MCMC) run at intervals of 100 generations until the MCMC run had reached convergence. Initial 500000 states (10%) of the sampled trees were discarded as burn-in. Three independent analyses starting from different random trees were conducted to assure that our analyses were not trapped in local optima and convergence of the MCMC run was reached from the independent random starting trees.

Bacterial strains and culture conditions.
P. putida KT2440[42] was grown in EM-medium or in minimal medium (MSB) supplemented with succinate as a sole carbon-source at 30°C. *E. coli* strains used in the study were grown in LB broth at 37°C. When necessary, kanamycin (50), was added in the concentration as indicated in parentheses (concentrations in µg/ml).

Growth of *P. putida* strains under iron starvation and illumination.
For growth experiments under iron starvation conditions, *P. putida* KT2440 was grown in a mineral salt medium (MSB) consisting of 6g/l K_2HPO_4, 3g/l KH_2PO_4 1g/l $(NH_4)_2SO_4$ and 0.2g/l $MgSO_4$ with succinate (4g/l) as sole carbon source. The pH of the medium was adjusted to pH 7.2. The cultures were diluted to OD600 = 0.01 of an overnight seed culture grown in the same medium. For each experiment six cultures of were grown under continuous blue, red and green light illumination ($30\mu E/m^2*s$), whereas additional six cultures were kept in the dark (wrapped in aluminum foil). Illumination was achieved in a custom-made dark incubator equipped with exchangeable LEDs (Luxeon Star, Lumileds-Phillips San Jose, CA, USA) to facilitate illumination of shaking liquid cultures. After 72 hours of incubation at 30°C, 120 rpm, the six identical cultures were combined and cells were pelleted by centrifugation. The pyoverdine containing supernatant was filtered by passing through a 0.22µM sterile filter unit.

Pyoverdine quantification.
Quantification of the pyoverdine content in culture supernatants was achieved fluorometrically by monitoring the intrinsic fluorescence of the pyoverdine siderophore. Pyoverdine emission spectra (410-600 nm) were recorded after excitation at 400nm using a Perkin-Elmer LS50B fluorescence spectrophotometer (Perkin-Elmer Ltd., Beaconsfield, England). In each independent experiment five measurements were carried out for each combined culture supernatant. For pyoverdine quantification the maximum of the emission peak (460 nm) was used. Values were normalized against the optical density (measured at 600 nm) of the corresponding cultures.

Biofilm quantification
A UV-sterilized 96-well microtiter plate (Greiner, Polystyrol), containing 100 µl MSB medium with succinate as sole carbon source in each well was inoculated with a starter culture of *P. putida* KT2440 grown overnight in the same medium. The plate was heat sealed with a light-transparent plastic foil and incubated for 72 hours at 30°C without shaking. Continuous blue-light illumination ($30\mu E/m^2*s$) was achieved using a custom-made 96-LED-module (Fa. Seltsam, Aachen, Germany) mounted perpendicular to the 96-well plate containing the *P. putida* KT2440 culture. A control was grown under identical conditions, but without illumination. Quantification of the biofilm adherent at the polystyrol walls of the single wells was achieved using a protocol adapted from O´Toole und Kolter (1998)[49]. Briefly, 20µl of a 1% crystal-violet solution (in 96% Ethanol) were added after the incubation to each of the 96 wells of the plate. The plate was incubated for 15 minutes at room-temperature. Planktonic cells were removed by washing carefully but thoroughly 10 times with distilled water. After drying of the crystal-violet stained biofilms - that now stick to the walls of the plate wells - 200µl of 96% Ethanol were added to each well and the absorbance of the dye was read at 650 nm using a SpectraMax-Plus 96-well plate reader (Molecular Devices, Sunnyvale, CA, USA).

Western blot analysis.
To facilitate the immunodetection of PpSB2-LOV in cell-free protein extracts of *P. putida KT2440*, the proteins were separated on SDS-polyacrylamide gels (16% acrylamide) and subsequently blotted onto PVDF membranes (BioRad Laboratories, Hercules, CA, USA). For the detection of PpSB2-LOV on the membrane, a specific polyclonal antiserum, raised against the purified protein was used (Eurogentec, Seraing, Belgium).

Cloning, heterologous expression and purification of PpSB1-LOV and PpSB2-LOV.
The PpSB1-LOV (Q88E39) encoding gene (PP4629) was PCR amplified by using the primer SB1_NdeI_fw: 5'-TCCAGTACATATGATCAACGCGCAATTGCTGC-3' and SB1_XhoI_rev: 5'-TACAATCTCGAGTCAGGCGCGTTCGTCGG-3'. PCR amplification was carried out using Turbo-Pfu™ DNA Polymerase according to the instructions of the manufacturer (Stratagene, LaJolla, CA, USA). The PCR amplified and *Nde*I, *Xho*I digested product was cloned into equally digested pET28a (Novagen, Merck, Darmstadt, Germany) furnishing PpSB1-LOV with an N-terminal hexa-histidine-tag resulting in the construct pET28a-PpSB1-LOV. Expression of PpSB2-LOV was achieved from a similar pET28a-PpSB2-LOV construct[28]. The recombinant protein was expressed in

E. coli BL21(DE3), employing an auto-induction medium adapted from Studier *et al.* (2005)[50], briefly, containing 12 g/l casein hydrolysate, 24 g/l yeast extract and 5 g/l glycerol in 100 mM, potassium phosphate buffer pH 7.0. To facilitate auto-induction of gene-expression the media was supplemented after sterilization of the media with 1% glucose (of a 50 g/l stock-solution) and 10% lactose (20 g/l stock-solution). Protein expression was carried out at 30°C for 30 hours. Immobilized Metal Affinity Chromatography (IMAC) purifications on Ni-NTA resin (QIAGEN, Hilden, Germany) were performed employing an ÄKTA Explorer FPLC system, as described previously[28].

Biochemical characterization: spectrophotometry, and HPLC separation of protein bound chromophore.

All protein samples were thoroughly dialyzed against 10 mM $NaPO_4$- buffer, 10 mM NaCl pH 8.0 using a 3000 Dalton cut-off Float-A-Lyser dialysis unit (Spectrum Laboratories Inc., Rancho Dominguez, CA,USA). Protein absorption spectra were recorded using a Beckmann DU-650 UV/VIS-spectrophotometer (Beckmann Coulter, Fullerton, CA, USA) whereas protein amounts were adjusted to an absorbance at 450 nm of approximately 0.2 to account for different apo-protein amounts in the preparations. Illumination of the protein samples with blue-light of 450 nm was achieved using a Led-Lenser®V8 lamp (Zweibrüder Optoelectronics, Soelingen-Germany) to accumulate the signaling state. As initial experiments indicated a very slow recovery of PpSB1-LOV, the dark recovery time of the protein had to be estimated by recording sequentially protein absorbance spectra after the protein was illuminated with blue-light.

Separation and quantification of protein bound FAD, FMN and Riboflavin was achieved by using a reverse phase C_{18}- HPLC column (250/4 Nucleodur, Macherey&Nagel, Düren, Germany, fitted with an 8/4 pre-column filled with the same material) connected to a Shimadzu LC10Ai HPLC system (Shimadzu Corporation, Kyoto, Japan). As mobile phases, 50 mM ammoniumacetate buffer pH 6.0 (Eluent B) and 70% acetonitrile in B (Eluent A) were used. As standards authentic FMN, FAD and Riboflavin purchased from FLUKA, Neu-Ulm, Germany were used.

To release the protein bound flavin chromophores, the samples were heat denatured (10 min, 99°C). The precipitated protein was removed by centrifugation at 4000 rpm, for 20 minutes. The flavin co-factor-containing supernatant was spun through a Microcon YM3 (Millipore, Billerica, MA, USA (MWCO 3000 Da) centrifugal concentrator device (13000 rpm, 1 hour) to remove any residual protein from the samples prior injection into the HPLC.

Elution was followed by absorption at 263 nm. The identity of the respective elution peak was verified by its typical flavin spectra. For identification and quantification, a calibration was performed using authentic compounds in a concentration range from 10 µM to 50 µM. Typical retention times for FAD, FMN and Riboflavin were 13,56 14,78 and 17,70 minutes, respectively.

HPLC protocol used for the separation of the different flavin species:

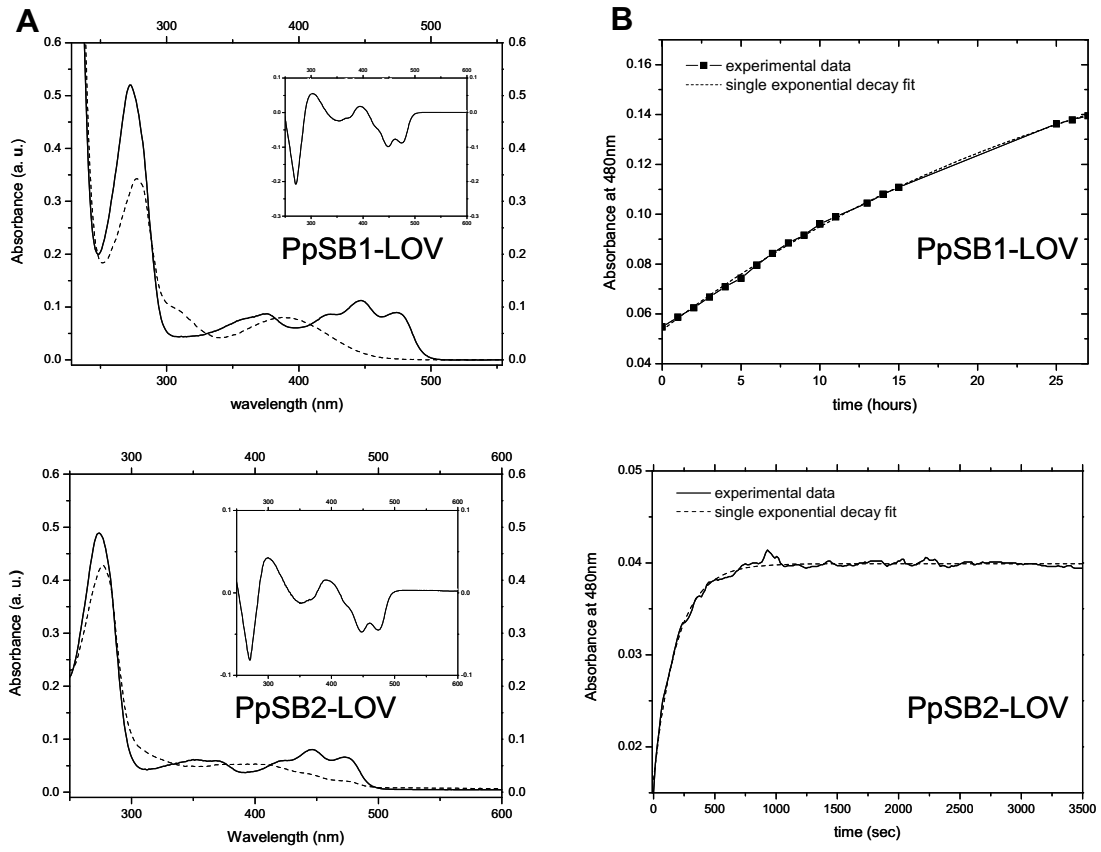
Time	Flow rate (ml/min)	% Eluent A	%Eluent B
0	0.63	5%	95%
20	0.63	40%	60%
22	0.8	40%	60%
29	0.63	5%	95%

B : 50mM NH_4Ac pH 6.0, A : 70% Acetonitrile in B

Results and Discussion

A second putative blue-light photoreceptor of the LOV family in *P. putida* KT2440

Previous BLAST analysis suggested the presence of two putative blue-light photoreceptor modules of the LOV family in *P. putida* KT2440[29]. In a previous study we could show that the protein PpSB2-LOV (Q88JB0) binds a flavin chromophore and shows photochemical activity similar to the blue-light sensitive signaling modules (LOV domains) in plant phototropin. A second putative LOV gene (named PpSB1-LOV in accordance with[28, 29], was PCR-amplified from *P. putida* KT2440 genomic DNA and subsequently cloned, into pET28a, similarly as described for the PpSB2-LOV gene. Sequence analysis demonstrated that PpSB1-LOV shares with PpSB2-LOV about 66% identical positions (78% semiconserved positions) on the amino acid (aa) level. The PpSB1-LOV protein was heterologously expressed as an N-terminal hexa-histidine fusion protein using *E. coli* BL21(DE3) as expression host. For comparison, Figure 1A shows the PpSB1-LOV and correspondingly the PpSB2-LOV dark state absorption spectra (solid line), as well the absorption spectra of the signaling state (dashed line) recorded after blue-light illumination of the sample. Due to a very slow recovery of PpSB1-LOV (revealed by initial experiments, data not shown), sequential absorption spectra for the protein over a period of 30 hours after an initial blue-light pulse had to be recorded in order to evaluate the dark recovery kinetics. The experimental data could be readily fitted using a single exponential decay curve (goodness of fit χ^2-test: R^2=0.999), thus revealing a very slow dark recovery time constant for PpSB1-LOV of τrec=29.5±1.5 hours (at 20°C). The corresponding recovery trace at 480 nm, extracted from the sequentially recorded absorption spectra is shown in Figure 1B. For comparison, the dark recovery time constant previously recorded for PpSB2-LOV was τrec=114 seconds (at 20°C) (Figure 1B). In order to evaluate the flavin-chromophore binding to PpSB1-LOV and PpSB2-LOV, with respect to the bound flavin species (FAD, FMN and riboflavin), a HPLC analysis was performed. The chromophores of purified PpSB1-LOV and PpSB2-LOV were extracted and subsequently separated using high-performance liquid chromatography (HPLC).



X

Figure 1: Comparative characterization of PpSB1-LOV and PpSB2-LOV. A) Blue-light sensitivity and photochemistry of PpSB1-LOV (upper half) and PpSB2-LOV (lower half). The Figures show the absorbance spectra of the dark state for both proteins (solid-line) and the corresponding absorbance spectra of the proteins after illumination with blue-light (dashed line). The insets show the light-dark difference spectra. **B)** Dark recovery kinetics for PpSB1-LOV and PpSB2-LOV, respectively. For PpSB1-LOV the dark recovery was monitored by recording sequential absorbance spectra over a period of 30 hours due to the very slow recovery of the protein. In the corresponding Figure, the kinetic trace at 480 nm was extracted from the sequential absorbance spectra. The experimental data could be well fitted using a single exponential decay curve. In case of PpSB2-LOV the dark recovery was monitored by recording the absorbance recovery at 480 nm over 3600 seconds. Correspondingly, the experimental data was fitted with a single exponential decay curve.

The assignment of the bound flavin species (FMN, FAD or riboflavin) was achieved by comparing the corresponding peak retentions time to those of authentic compounds (Table 1). The HPLC analysis of the protein-bound flavin content revealed that PpSB1-LOV binds only (100%) FMN, but no detecTable other flavin species. Contrarily, PpSB2-LOV seems to contain a mixture of FMN (62%) and riboflavin (38%).

Table 1: Comparison of flavin-chromophore content of PpSB1-LOV and PpSB2-LOV as derived from high performance liquid chromatography (HPLC).

	FAD	**FMN**	**Riboflavin**
PpSB1-LOV	n.d	100%	n.d
PpSB2-LOV	n.d	62%	38%

n.d.: not detecTable under the used conditions (see experimental section for details)

Structural determinants of the dark recovery kinetics.

The pressing question that arises is what structurally determines the dramatically different dark recovery kinetics for two proteins that on the sequence level share about 66% of identical aa positions. Although, the mechanistic basis of the primary events in the photocycle of LOV domains are extensively studied and largely understood, the mechanisms that govern the dark recovery reaction on the other hand remain largely elusive. Nevertheless, several hypotheses were recently brought forward:

In a recent study Kennis, Alexandre and co-workers[51] suggested, in accordance with a previous hypothesis of Swartz *et al.*[24], a base catalyzed recovery reaction. Mechanistically it was proposed that in a base-catalyzed back reaction, a proton has to be abstracted from N(5) of the flavin isoalloxazine ring. The breaking of the Cys-S-C(4a) and the regeneration of the C(4a)-N(5) double bond probably occurs spontaneously, to relieve strain in the protein and the flavin ring system caused by the proton abstraction from N(5)[24]. As proton-abstracting groups, two surface exposed histidine residues in the *Avena sativa* phot1- LOV2 domain were suggested[51]. Although, they are located in a distance of about 12 Å from the flavin ring, base catalysis was suggested to occur via a postulated hydrogen bond network of the base, chromophore and intraprotein water molecules[51].

For PpSB1-LOV (characterized by a very slow dark recovery) and PpSB2-LOV (characterized by a fast recovery), the presence of four histidines in PpSB2-LOV and correspondingly the complete absence of

histidine residues in PpSB1-LOV might indeed imply a similar mechanism. In the PpSB1-LOV / PpSB2-LOV system this explanation seems feasible; especially as in PpSB2-LOV a histidine residue (His64) is probably located in close proximity to the flavin isoalloxazine ring system (structural information derived from the homology model of PpSB2-LOV described by Krauss *et al.* (2005)[28]). In PpSB1-LOV the same position is occupied by an arginine residue (Arg61). Thus, the slow dark recovery of PpSB1-LOV might be due to the absence of a strong proton acceptor in the vicinity of the flavin-ring system and *vice versa* the fast recovery of PpSB2-LOV might be driven correspondingly from the above mentioned histidine (His64) as proton-accepting group.

Another hypothesis based on a random mutagenesis study on the *Avena sativa*(As) phot1-LOV2 domain, suggested the involvement of an isoleucine residue located within van der Waals distance from the highly conserved photoactive cysteine. The isoleucine probably makes contact via its methyl-group to the sulfhydryl group of the photoactive cysteine residue[52], causing steric stabilization of the C(4a)-thiol(S) bond in the LOV signaling state. Mutation of this isoleucine to valine or leucine resulted in a dark recovery that was faster compared to the wildtype protein. The authors consequently argued the mutation to valine or leucine alters the native, light-driven support of the C(4a)-S bond and thus enables a faster dark recovery[52].

In PpSB1-LOV and PpSB2-LOV the sequence region of the isoleucine residue identified in Asphot1-LOV2 as hot-spot that strongly influences the dark recovery reaction, is conserved between the two proteins. In both proteins the position is occupied by a leucine, which might in accordance to the above hypothesis account for the relatively fast recovery reaction in PpSB2-LOV. However, quite to the contrary, the same argument fails to explain the slow recovery observed for PpSB1-LOV.

Therefore, as in particular the protein region around the leucine residue is highly conserved in both of the paralogous *Pseudomonas* proteins, the dramatic difference between the two cannot be explained by a steric stabilization effect as in the case of the Asphot1-LOV2 domain.

This implies the involvement of additional mechanisms in the PpSB1-LOV and PpSB2-LOV recovery that result in the observed differences and hence structurally distinguish the two. The high degree of sequence conservation between PpSB1-LOV and PpSB2-LOV (66% identical positions) and the concomitant dramatic difference in their dark recovery time constants, should render the pair of paralogous proteins PpSB1-LOV and PpSB2-LOV a perfect model system to study the mechanisms of the dark recovery process, by means of site-directed or random mutagenesis.

Presence of conserved LOV signaling modules arranged in conserved gene clusters suggest functional convergent responses in saprotrophic Pseudomonads.

A BLAST-analysis was performed to identify PpSB1-LOV and PpSB2-LOV homologous protein sequences in the genome of other *Pseudomonas* species. Only those LOV sequences were considered as hits, in which the LOV domain- specific canonical sequence motif GXNCRFLQG was changed to YQDCRFLQG, a feature that clearly distinguishes the *P. putida* LOV proteins from all other so far identified plant and bacterial LOV domains. Using this criterion, homologous sequences could be identified in *P. putida* W619, *P. putida* F1, *P. putida* GB1, *P. fluorescence* Pf-5, *P. fluorescens* PfO1 as well as in *P. mendocina* YMP and in *P. stutzeri* A1501. Whereas, the first two *P. putida* strains W619 and F1 as well as the previously analyzed *P .putida* KT2440 possess two such LOV proteins encoded in different locations in their genome, the latter five *Pseudomonas* strains encode only one LOV protein. The multiple sequence alignment of those sequences is provided in supplementary Figure 1. Pairwise identities to PpSB1-LOV and PpSB2-LOV are shown in Table 2.

Table 2: Pairwise sequence identity of the newly identified Pseudomonas LOV homologs.

LOV protein	identical positions (%)	assignment[$]	identical positions (%)
	PpSB1-LOV		**PpSB2-LOV**
PpSB1-LOV	100		66
PpSB2-LOV	66		100
P.putida W619-1	86	←	64
P.putida W619-2	61	→	83
P.putida F1-1	98	←	65
P.putida F1-2	65	←	98
P.putida GB1	92	←	64
P.fluorescens Pf-5	65	≈	64
P.fluorescens PfO-1	55	≈	55
P.mendocina YMP	64	≈	70
P.stutzeri A1501	60	≈	68

[$]: assignment to one or the other group of sensory box protein sequences is indicated by an arrow. In case sequence identity comparison allowed an unequivocal assignment to a particular sensory box protein group (PpSB1-like or PpSB2-like), the corresponding value is highlighted in grey. If no unequivocal assignment was possible an ≈ is shown in the assignment column.

Interestingly, for the first three newly identified *P. putida* strains a clear differentiation into PpSB1-LOV-like and PpSB2-LOV-like was possible, based on higher sequence identity to either PpSB1-LOV or PpSB2-LOV. However, for the other *Pseudomonas* strains, such an assignment was not possible, due to a less pronounced affinity for one or the group of sequences (see Table 2).

Therefore, to gain some insight into the evolutionary history of the two LOV signaling modules in *Pseudomonads*, we reconstructed a phylogenetic tree based amino acid sequence data (Figure 2). Additionally, LOV protein genomic loci, or functionally conserved proteins that flank the corresponding LOV protein(s) in the different *Pseudomonas* strains were compared and plotted to the phylogentic tree in Figure 2. Both Next-Neighbour Interchange (NNI) and Subtree Pruning and Recrafting SRP) algorithms as well as Bayesian posterior probability calculations and single Important Quartet Puzzling and Next Neighbour Interchange (IQPNNI) runs result in the same tree topology depicted in Figure 2. The phylogenetic analysis thus suggests the two *P. fluorescens* LOV sequences as ancestral to the *P. putida* LOV protein

sequences, with a very recent duplication event giving rise to PpSB1-LOV-like and PpSB2-LOV-like sequences in all *P. putida* strains. This duplication might have occurred after the speciation into *P. putida* and *P. fluorescens;* or alternatively it implies an earlier duplication event with subsequent loss of one LOV protein module in the two *P. fluorescens* strains, as well as in two non-fluorescent *Pseudomonas* species (*P. stutzeri and P. mendocina*).

Conserved operon-like structures in all LOV domain harbouring Pseudomonads suggest a common light-signal transduction mode

Interestingly, the genome context in which the respective PpSB1-LOV and PpSB2-LOV homologs are found, is more or less conserved throughout the different *P. putida* strains (Figure 2). For example a MerR-like[53] transcriptional regulator comprising a helix-turn-helix (HTH) DNA binding motif is found in nearly all *Pseudomonas* strains in close proximity to the corresponding LOV protein(s). Furthermore, highly conserved in several *Pseudomonas* strains are LysR-like (HTH) transcriptional regulators[54].

This co-localization of isolated LOV signaling modules (PpSB1-LOV/PpSB2-LOV) that do not contain a fused effector domain, together with transcriptional regulators of the HTH-family (MerR- or LysR-like)[53, 55] in all analyzed *Pseudomonas* strains, implies a mode for the signal transduction from the light sensor to the downstream targets that may initiate a light-driven response. Therefore, based on our bioinformatic analysis, we suggest the HTH- transcriptional regulators as the LOV associated, but not fused, effector domains in *P. putida* KT2440 and probably in other Pseudomonads. This suggests a light driven initiation or regulation of the expression of certain genes in response to a blue-light stimulus. Moreover, although an interaction of a LOV domain with a transcriptional regulator, has so far not been proven experimentally, comparative sequence analyses of putative LOV photosensors show that a number of proteins exist in nature in which a LOV domain is directly fused to a HTH transcriptional regulator.

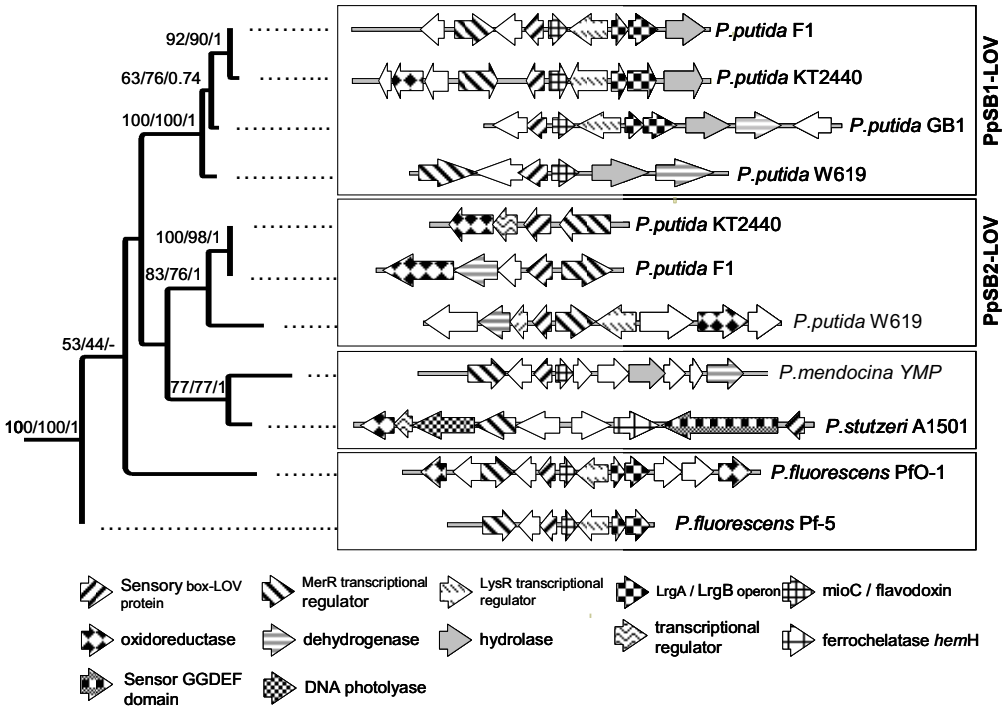


Figure 2: Phylogenetic tree reconstructed for the different short LOV proteins found in several *Pseudomonas* species. Conserved Operon structures (or classes of functionally conserved proteins) are shown at the particular branch of the tree. The bootstrapped maximum-likelihood tree was constructed using PhyML and MultiPhyl-Online Server by applying both Next Neighbour Interchange (NNI) and Subtree Pruning and Recrafting (SRP) algorithms for the heuristic tree search. Furthermore, BEAST v1.4.5 was used to infer Bayesian posterior probabilities for each bifurcation of the ML-tree. Support values of 100 bootstrap replicates of the NNI and SRP calculation as well as BEAST posterior probability values (in that order) are depicted at the corresponding branch. Only values higher than 50% are shown

For example the LOV proteins found in *Erythrobacter litoralis* (Q2NB98), *Novosphingobium aromaticivorans* (Q2G8Z7), *Sphingopyxis alaskensis* (Q1GUF5), *Sphingomonas sp.* SKA58 (Q1N7J1) and *Thiomicrospira denitrificans* (Q30NS0) all possess a LOV domain fused to a HTH DNA-binding domain, thus corroborating at least the possibility of a signal-transduction process between the two partners with the LOV domain acting as the sensor- and the HTH-transcriptional regulator as the effector- or output domain.

Implications of the genome proximity for the biological role of the LOV proteins in *P. putida* KT2440

Another highly conserved feature of the PpSB1-LOV protein containing genomic region that distinguishes the

PpSB1-LOV locus from the PpSB2-LOV locus, is the presence of an LrgA/LrgB operon-like structure within the PpSB1-LOV genomic region. The LrgA/LrgB system together with its homologs cidA/cidB were identified to be involved in *Staphylococcus aureaus* mediated cell death and lysis[56] and were furthermore suggested to play a role in biofilm development and adherence in this organism[57].

Moreover, the LrgA/LrgB operon is also present in the two probably ancestral *P. fluorescens* strains, thus on the one hand corroborating our hypothesis on the ancestry of the PpSB1-LOV-like systems over the PpSB2-LOV-like systems and on the other hand pointing toward biological processes that might be influenced by one of the two light sensors.
A more extensive analysis of the wider genomic proximity of the *P. putida* KT2440 LOV-proteins revealed a clustering of proteins that are annotated to play a role in the iron-starvation response in *P. putida* KT2440 (derived from genome region comparison using tools implemented in the PutidaList server, data not shown). Especially, the genomic proximity of the PpSB1-LOV protein is rich in siderophore-receptors, outer-membrane ferric-iron and dicitrate receptors, as well as extracytoplasmatic function (ECF)-Sigma Factors. The latter ones are implied in the control of transcription in response to external stimuli (that might include iron-starvation). Whereas, the different receptors (e.g. siderophore-uptake and iron-uptake), might be involved in regulating cellular iron-homeostasis and thus counteracting iron-limitation by facilitating ferric-iron uptake.
Taken together, the presence of iron-starvation regulated genes like siderophore receptors, iron-transporters as well as the presence of the LrgA/B operon implied in the iron-starvation response and biofilm development (a process which is intimately linked to iron-acquisition in Pseudomonads) in close genomic proximity to the two LOV-photosensors in *P. putida* KT2440 might point toward a biological role for light driven responses in this organism. Furthermore, the co-localization of HTH-transcriptional regulators suggests a mode for the signal transduction chain: the blue-light stimulus that is received in the LOV domain might subsequently be relayed to the HTH-transcriptional regulator (via induction of protein-protein interaction or release), which in turn might activate the transcription of certain genes that enable the organisms to cope with the effects caused by the blue-light stimulus.

The central question for *P. putida* and likewise for the majority of the other bacterial species that possess a LOV signaling module remains: what is the biological relevance of a blue-light sensor in those organisms, and moreover, what is the biological reason why certain bacteria should be able to perceive a blue-light stimulus whereas others a completely blind towards it.

Blue-light effects on *P. putida* KT2440

In accordance with the genome proximity analysis, we attempted to address the question of the biological relevance of a blue-light stimulus for *P. putida* KT2440. Therefore, we studied a possible blue-light dependency of certain iron-deficiency dependent phenomena in the wildtype strain, e.g. pyoverdine-siderophore production[58] and biofilm formation.

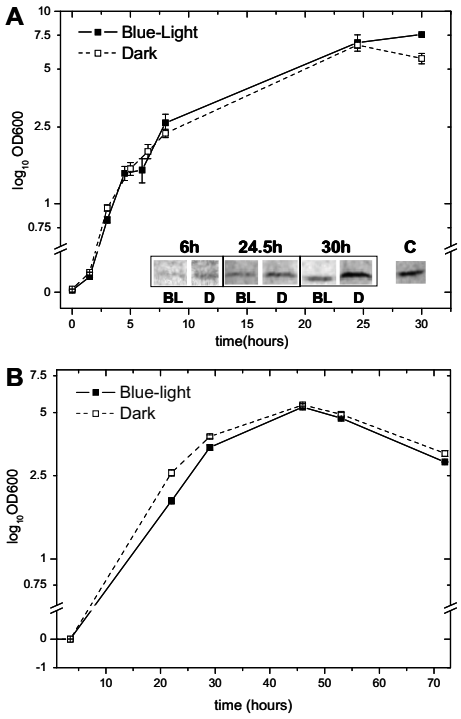


Figure 3: Growth behaviour of *P. putida* KT2440 under blue-light and in the dark. Microbial growth was monitored photometrically by measuring the OD600 (optical density measured at 600nm) of the growing cultures. **A)** The strain was grown in full-media (iron-rich conditions), see experimental section for details. Whereas, three identical cultures were grown under continuous blue-light illumination and six cultures were grown in the dark under identical conditions. The inset depicts the accumulation of the PpSB2-LOV protein in the cells under blue-light (BL) and correspondingly in the dark (D). PpSB2-LOV detection was carried out immunologically as described in the experimental section. **B)** *P. putida* KT2440 was grown in minimal media lacking an external iron source (MSB). Three identical cultures were grown under continuous blue-light illumination whereas three cultures were kept in the dark. In both Figures, the values represent the mean of the OD600 measurement from corresponding identical cultures. Error bars derive from the associated standard deviation.

Growth-phase dependent PpSB2-LOV accumulation in *P. putida* KT2440

The wildtype strain, *P. putida* KT2440 was grown in shaking liquid cultures at 30°C in i) iron-rich full-media (EM) (Figure 3A) and in ii) a minimal-medium supplemented with succinate as sole carbon source (MSB) that lacked externally added iron (Figure 3B). Therefore, growth in the minimal-media should induce severe iron-starvation in the strain. Several identical-cultures were continuously illuminated with blue-light, whereas the same number of cultures was kept in the dark (covered with aluminum-foil). During growth in iron-rich media, samples were taken at the time-points 6 hours, 24.5 hours and 30 hours after inoculation. Cells were lysed and the crude cell free extracts were separated using SDS-PAGE and subsequently blotted to a PVDF membrane.

The PpSB2-LOV protein in cell-free extracts of *P. putida* KT2440 could be detected immunologically using a specific polyclonal anti-serum raised against purified PpSB2-LOV. The accumulation of PpSB2-LOV during the growth of the strain in blue-light and in the dark is shown in Figure 4A as an inset.
The western-blot suggests a growth phase dependent accumulation of PpSB2-LOV, when grown in full-media, thus, suggesting a cell-density dependent expression of PpSB2-LOV in *P. putida* KT2440. It should be noted here, that the antiserum used to detect PpSB2-LOV shows a cross-reactivity against PpSB1-LOV due to the high degree of sequence conservation between the two proteins. Although cross-reactivity towards PpSB1-LOV is weak, the signal detected by the western analysis has to be considered as the sum of the two LOV proteins in the cell. This effect might especially be relevant in a scenario in which a large PpSB1-LOV amount might overcompensate for a small amount of PpSB2-LOV. As this can not be ruled out without specifically probing the PpSB1-LOV accumulation by terms of a separate western analysis using a PpSB1-LOV antiserum, the result has to be considered with care.

Light dependency of the pyoverdine production in *P.putida* KT2440

To monitor the levels of the fluorescent iron-siderophore pyoverdine that is only produced in response to iron-limiting conditions, *P. putida* KT2440 was grown in shaking liquid cultures (MSB media) for 72 hours (late stationary phase).
Several identical cultures were continuously illuminated with blue-light, red-light and green-light of equal intensity. Likewise, several identical cultures were kept in the dark as described before. After 72 hours of growth, equally treated cell-free culture supernantants from light and dark grown cultures were used to assay the pyoverdine levels, respectively.
The amount of pyoverdine in the supernatants was estimated fluorimetrically and the values were normalized against the optical-density of the corresponding culture. Figure 4A shows the comparison of estimated pyoverdine amounts that are produced in cultures grown under blue-light, red-light, green-light and in the dark. The data shows an approximate 2-fold increase in the production of the iron-siderophore pyoverdine in cultures of *P. putida* KT2440 that are grown under continuous blue-light illumination compared to cultures that were kept in the dark. Furthermore, no increase in pyoverdine production could be detected in cultures grown under continuous red- and green-light illumination. The inset in Figure 4A shows the immunological detection of PpSB2-LOV. Western blot analysis suggests a constitutive expression of PpSB2-LOV under all studied light conditions.
Taken together, the exclusive increase of the iron-siderophore pyoverdine levels in blue-light grown cultures supernatants of *P. putida* KT2440, strongly suggests the action of a blue-light photoreceptor in this process. However, the involvement of the two LOV-family photoreceptors has yet to be proven by molecular methods.

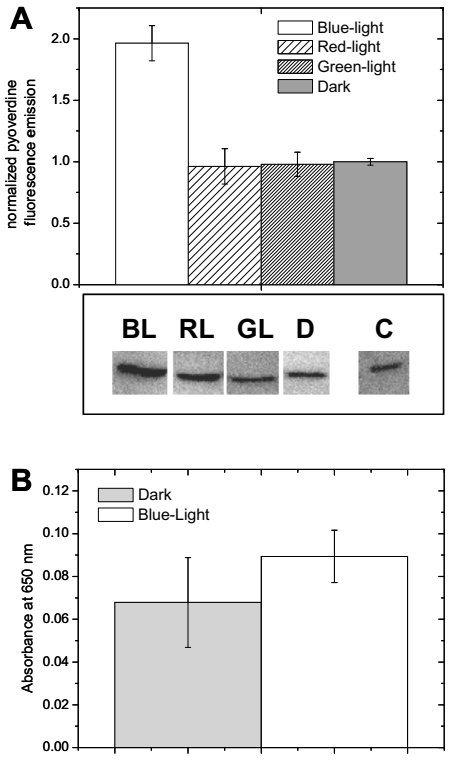


Figure 4: Summary of the observed blue-light effects in *P.putida* KT2440. A) Blue-light dependency of the pyoverdine production in *P.putida* KT2440. Average normalized pyoverdine fluorescence emission intensities of *P. putida* KT2440 culture supernatants grown under iron starvation in different light conditions. The red light (RL, wide diagonally hatched bar), green light (GL, narrow diagonally hatched bar) and blue light (BL, white bar) intensities were normalized against the optical density of the corresponding cultures. The pyoverdine fluorescence emission intensity of dark grown cell supernatants (D, gray bar) was arbitrarily normalized to 100% (1.0). The values represent the mean of four independent identical experiments with seven identical cultures grown under blue-light (BL) and correspondingly seven cultures that were kept in the dark (D). For GL and RL experiments two independent sets of experiments with seven identical cultures each, were conducted. Error bars represent the standard deviation of the multiple experiments. The accretion to Figure 3A, depicts the PpSB2-LOV protein accumulation probed by western analysis as described in the experimental section**. B)** Blue-light dependency of the *P .putida* KT2440 biofilm formation. The assay was conducted as described in the experimental section. Values represent the mean of three independent identical experiments performed under blue-light illumination (white bar) and in the dark (grey bar). The error bars represent the associated standard deviation of the identical experiments.

Dependency of the biofilm formation in *P. putida* KT2440 on blue-light

An initial characterization of the blue-light dependency of the biofilm formation of *P. putida* KT2440 was conducted in a static biofilm assay as described in the methods section. Biofilms were grown under iron-limiting conditions with and without continuous blue-light illumination. Figure 4C shows
the comparison of the biofilm-biomass (expressed in terms of absorbance at 650 nm, of the cell-staining crystal-violet dye, that was used for quantification). Although the associated experimental error of the used method is quite large, three independent experiments confirmed the same overall behaviour, namely a slight induction of biofilm formation by blue-light. It has to be noted that this characterization is preliminary since

more sophisticated methods have to be employed to prove the blue-light effect, like confocal microscopy of static and flow-cell grown biofilms.

A possible physiological scenario for the blue-light dependency of the iron-starvation response in *P. putida* KT2440

The question central to the discussion of blue-light effects in non-photosynthetic bacteria is, whether or not blue-light is able to effect (either beneficially or detrimentally) a bacterial system. Interestingly, the answer might come from the best biochemically and genetically characterized microbe *Escherichia coli.* Already in 1991, Miyamoto and co-workers demonstrated a visible light phenotype for *E. coli* K12[59]. When a gene designated *vis*A was inactivated, the resulting mutants were sensitive towards visible light or in other words illumination resulted in an efficient killing of *vis*A mutant bacteria. The authors identified the corresponding gene as a ferrochelatase (probably identical to the previously designated *hemH* gene in *E. coli* K12). Ferrochelatase (protoheme ferrolyase, EC 4.99.1.1), that is ubiquitous from bacteria to man, is catalyzing the final step of the heme biosynthesis, namely the incorporation of ferrous iron into protoporphyrin IX (PPIX) ring system[60]. This process represents one bottleneck in the heme biosynthesis pathway since unavailability of iron will impair heme synthesis at this reaction step of the pathway and thus result ultimately in PPIX accumulation. Consequently, *hemH* was later proposed to be subject to regulation by iron availability[61]. Furthermore Baysse, et al. (2003) suggested that PPIX accumulation in a strain of *P. fluorescens* can cause an *E.coli* K12-like *vis*A-phenotype[62]. The phototoxicity of PPIX is probably related to the light-dependent formation of reactive oxygen species (ROS) from the porphyrin ring system of PPIX. Mechanistically this is driven by the excitation of accumulated PPIX by blue-light, that eventually results in the formation of free radicals and thus ROS[3-5] induced oxidative stress/damage to the cell. Therefore, it seems reasonable for the cell to adapt to iron-limitation and additionally to the presence of blue-light as environmental stimulus. Several of the observed blue-light dependencies in the *P. putida* KT2440 iron-starvation response that were discussed above might participate to counteract the detrimental effects of ROS and thus of blue-light. The iron-siderophore pyoverdine is responsible for efficient scavenging of even trace amounts of ferrous iron from the environment, therefore feeding into one bottleneck of the heme biosynthesis pathway, namely the incorporation of iron into the protoporphyrin IX ring. Thus, iron-limitation alone is activating the pyoverdine production, whereas blue-light might additionally (up)-regulate the process. Furthermore, the blue-light dependent increase in biofilm formation might be another method utilized by cells to overcome the direct exposure of the microorganism to the harmful blue-light irradiation.
Elaborating on this hypothesis, cells buried within the biofilm will be less susceptible to the radiation effects and thus biofilm formation might increase survivability of the strain under blue-light irradiation.

Conclusions

With this study we demonstrated the presence of two LOV-photosensor modules in *P. putida* KT2440. Those two signaling modules, although very similar in sequence (66% identical aa) and probably structure, show dramatically different dark recovery kinetics and a variability in the chromophore specifity when expressed in the heterologous host as revealed by HPLC analysis of the chromophore content. A bioinformatic analysis suggested the involvement of the two LOV-proteins in responses to iron-limitation, and moreover implied a HTH-transcritional regulator as LOV associated, but not fused effector in the signaling cascade. Furthermore, based on initial photo-physiological experiments using the *P. putida* KT2440 wildtype strain, we could demonstrate the dependency of several iron-limitation regulated responses, such as pyoverdine siderophore production and biofilm formation on blue-light. This strongly implies the involvement of a blue-light photoreceptor in this process, especially since in case of the pyoverdine production effect no response could be detected for red- and green-light.
We elaborate a possible route in which blue-light might influence or harm *P. putida* KT2440. As comparative genome mining demonstrates the conservation of the LOV-photoreceptor modules found in *P. putida* KT2440 also in other fluorescent and non-fluorescent Pseudomonads, similar blue-light responses might be of importance among the group of saprotrophic Pseudomonads in general.
This study provides first hints on a possible biological role of blue-light in Pseudomonads and suggests the conservation of the LOV paradigm in this taxonomical lineage. Nevertheless, molecular studies, inactivating independently both LOV genes in *P. putida* KT2440, should follow, and are currently in progress in our laboratory, to unequivocally prove the involvement of the LOV photoreceptor(s) in the mention responses.

Acknowledgements

The authors would like to thank Astrid Wirtz for her excellent help with the HPLC analysis and Franco Circolone (both at the Institut für Molekulare Enzymtechnologie, Juelich, Germany) for his assistance in purifying and characterizing PpSB1-LOV. This work was supported by grants from the Deutsche Forschungsgemeinschaft (DFG) within the Forschergruppe FOR526 “Sensory blue light photoreceptors”.

References

1. A. Bekker, H. D. Holland, P. L. Wang, D. Rumble, 3rd, H. J. Stein, J. L. Hannah, L. L. Coetzee and N. J. Beukes, Dating the rise of atmospheric oxygen, *Nature*, 2004, **427**, 117-120.
2. D. I. Pattison and M. J. Davies, Actions of ultraviolet light on cellular structures, *Exs*, 2006, 131-157.
3. S. W. Ryter, H. P. Kim, A. Hoetzel, J. W. Park, K. Nakahira, X. Wang and A. M. Choi, Mechanisms of cell death in oxidative stress, *Antioxid Redox Signal*, 2007, **9**, 49-89.

4. J. D. Spikes, Photosensitization., in *The Science of Photobiology*, ed. K. C. Smith, Plenum Press, New York and London., 1989, pp. 79-110.
5. E. Clo, J. W. Snyder, P. R. Ogilby and K. V. Gothelf, Control and selectivity of photosensitized singlet oxygen production: challenges in complex biological systems, *Chembiochem*, 2007, **8**, 475-481.
6. J. M. Christie, Phototropin blue-light receptors, *Annu Rev Plant Biol*, 2007, **58**, 21-45.
7. D. Bhaya, Light matters: phototaxis and signal transduction in unicellular cyanobacteria, *Mol Microbiol*, 2004, **53**, 745-754.
8. L. O. Essen and T. Klar, Light-driven DNA repair by photolyases, *Cell Mol Life Sci*, 2006, **63**, 1266-1277.
9. J. Hughes, T. Lamparter, F. Mittmann, E. Hartmann, W. Gartner, A. Wilde and T. Borner, A prokaryotic phytochrome, *Nature*, 1997, **386**, 663.
10. H. Smith, Phytochrome and photomorphogenesis in plants, *Nature*, 1970, **227**, 665-668.
11. K. H. Jung, V. D. Trivedi and J. L. Spudich, Demonstration of a sensory rhodopsin in eubacteria, *Mol Microbiol*, 2003, **47**, 1513-1522.
12. O. A. Sineshchekov, K. H. Jung and J. L. Spudich, Two rhodopsins mediate phototaxis to low- and high-intensity light in Chlamydomonas reinhardtii, *Proc Natl Acad Sci U S A*, 2002, **99**, 8689-8694.
13. R. Kort, W. D. Hoff, M. Van West, A. R. Kroon, S. M. Hoffer, K. H. Vlieg, W. Crielaand, J. J. Van Beeumen and K. J. Hellingwerf, The xanthopsins: a new family of eubacterial blue-light photoreceptors, *Embo J*, 1996, **15**, 3209-3218.
14. M. Ahmad and A. R. Cashmore, HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor, *Nature*, 1993, **366**, 162-166.
15. M. Gomelsky and G. Klug, BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms, *Trends Biochem Sci*, 2002, **27**, 497-500.
16. A. Losi, The bacterial counterparts of plant phototropins, *Photochem Photobiol Sci*, 2004, **3**, 566-574.
17. S. Crosson, S. Rajagopal and K. Moffat, The LOV domain family: photoresponsive signaling modules coupled to diverse output domains, *Biochemistry*, 2003, **42**, 2-10.
18. K. Huang, T. Merkle and C. Beck, Isolation and characterization of a Chlamydomonas gene that encodes a putative blue-light photoreceptor of the phototropin family. , *Physiol. Plant* 2002, **115**, 613-622.
19. W. R. Briggs and J. M. Christie, Phototropins 1 and 2: versatile plant blue-light receptors, *Trends Plant Sci*, 2002, **7**, 204-210.
20. M. Ohgishi, K. Saji, K. Okada and T. Sakai, Functional analysis of each blue light receptor, cry1, cry2, phot1, and phot2, by using combinatorial multiple mutants in Arabidopsis, *Proc Natl Acad Sci U S A*, 2004, **101**, 2223-2228.
21. K. Huang and C. F. Beck, Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga Chlamydomonas reinhardtii, *Proc Natl Acad Sci U S A*, 2003, **100**, 6269-6274.
22. J. T. Kennis, S. Crosson, M. Gauden, I. H. van Stokkum, K. Moffat and R. van Grondelle, Primary reactions of the LOV2 domain of phototropin, a plant blue-light photoreceptor, *Biochemistry*, 2003, **42**, 3385-3392.
23. T. Kottke, J. Heberle, D. Hehn, B. Dick and P. Hegemann, Phot-LOV1: photocycle of a blue-light receptor domain from the green alga Chlamydomonas reinhardtii, *Biophys J*, 2003, **84**, 1192-1201.
24. T. E. Swartz, S. B. Corchnoy, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs and R. A. Bogomolni, The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin, *J Biol Chem*, 2001, **276**, 36493-36500.
25. S. Crosson and K. Moffat, Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction, *Proc Natl Acad Sci U S A*, 2001, **98**, 2995-3000.
26. S. Crosson and K. Moffat, Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch, *Plant Cell*, 2002, **14**, 1067-1075.
27. R. Fedorov, I. Schlichting, E. Hartmann, T. Domratcheva, M. Fuhrmann and P. Hegemann, Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii, *Biophys J*, 2003, **84**, 2474-2482.
28. U. Krauss, A. Losi, W. Gartner, K. E. Jaeger and T. Eggert, Initial characterization of a blue-light sensing, phototropin-related protein from Pseudomonas putida: a paradigm for an extended LOV construct, *Phys Chem Chem Phys*, 2005, **7**, 2804-2811.
29. A. Losi, The bacterial counterparts of plants phototropins,, *Photochem Photobiol Sci*, 2004, **3**, 566-574.
30. A. Losi, E. Polverini, B. Quest and W. Gartner, First evidence for phototropin-related blue-light receptors in prokaryotes, *Biophys J*, 2002, **82**, 2627-2634.
31. K. Zikihara, T. Iwata, D. Matsuoka, H. Kandori, T. Todo and S. Tokutomi, Photoreaction cycle of the light, oxygen, and voltage domain in FKF1 determined by low-temperature absorption spectroscopy, *Biochemistry*, 2006, **45**, 10828-10837.
32. B. D. Zoltowski, C. Schwerdtfeger, J. Widom, J. J. Loros, A. M. Bilwes, J. C. Dunlap and B. R. Crane, Conformational switching in the fungal light sensor Vivid, *Science*, 2007, **316**, 1054-1057.
33. R. Narikawa, K. Zikihara, K. Okajima, Y. Ochiai, M. Katayama, Y. Shichida, S. Tokutomi and M. Ikeuchi, Three putative photosensory light, oxygen or voltage (LOV) domains with distinct biochemical properties from the filamentous cyanobacterium Anabaena sp. PCC 7120, *Photochem Photobiol*, 2006, **82**, 1627-1633.
34. Z. Cao, Buttani, V., Losi, A., and Gärtner, W., A blue light inducible two component signal transduction system in
the plant pathogen Pseudomonas syringae pv. tomato, *submitted for publication*, 2007.
35. T. E. Swartz, T. S. Tseng, M. A. Frederickson, G. Paris, D. J. Comerci, G. Rajashekara, J. G. Kim, M. B. Mudgett, G. A. Splitter, R. A. Ugalde, F. A. Goldbaum, W. R. Briggs and R. A. Bogomolni, Blue-light-activated histidine kinases: two-component sensors in bacteria, *Science*, 2007, **317**, 1090-1093.
36. T. A. Gaidenko, T. J. Kim, A. L. Weigel, M. S. Brody and C. W. Price, The blue-light receptor YtvA acts in the environmental stress signaling pathway of Bacillus subtilis, *J Bacteriol*, 2006, **188**, 6387-6395.
37. N. Suzuki, N. Takaya, T. Hoshino and A. Nakamura, Enhancement of a sigma(B)-dependent stress response in Bacillus subtilis by light via YtvA photoreceptor, *J Gen Appl Microbiol*, 2007, **53**, 81-88.
38. M. Avila-Perez, K. J. Hellingwerf and R. Kort, Blue light activates the sigmaB-dependent stress response of Bacillus subtilis via YtvA, *J Bacteriol*, 2006, **188**, 6411-6414.

39. D. E. Somers, T. F. Schultz, M. Milnamow and S. A. Kay, ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis, *Cell*, 2000, **101**, 319-329.
40. W. Y. Kim, S. Fujiwara, S. S. Suh, J. Kim, Y. Kim, L. Han, K. David, J. Putterill, H. G. Nam and D. E. Somers, ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light, *Nature*, 2007, **449**, 356-360.
41. B. Thomas, Light signals and flowering, *J Exp Bot*, 2006, **57**, 3387-3393.
42. K. E. Nelson, C. Weinel, I. T. Paulsen, R. J. Dodson, H. Hilbert, V. A. Martins dos Santos, D. E. Fouts, S. R. Gill, M. Pop, M. Holmes, L. Brinkac, M. Beanan, R. T. DeBoy, S. Daugherty, J. Kolonay, R. Madupu, W. Nelson, O. White, J. Peterson, H. Khouri, I. Hance, P. Chris Lee, E. Holtzapple, D. Scanlan, K. Tran, A. Moazzez, T. Utterback, M. Rizzo, K. Lee, D. Kosack, D. Moestl, H. Wedler, J. Lauber, D. Stjepandic, J. Hoheisel, M. Straetz, S. Heim, C. Kiewitz, J. A. Eisen, K. N. Timmis, A. Dusterhoft, B. Tummler and C. M. Fraser, Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440, *Environ Microbiol*, 2002, **4**, 799-808.
43. K. Nicholas, Nicholas, H.J., and Deerfield, D. , GeneDoc: Analysis and Visualization of Genetic Variation., *EMBNEWS*, 1997 **4**, 14.
44. S. Guindon, F. Lethiec, P. Duroux and O. Gascuel, PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference, *Nucleic Acids Res*, 2005, **33**, W557-559.
45. T. M. Keane, T. J. Naughton and J. O. McInerney, MultiPhyl: a high-throughput phylogenomics webserver using distributed computing, *Nucleic Acids Res*, 2007, **35**, W33-37.
46. S. Vinh le and A. Von Haeseler, IQPNNI: moving fast through tree space and stopping in time, *Mol Biol Evol*, 2004, **21**, 1565-1571.
47. A. J. Drummond, Rambaut, A.,, BEAST v1.4, Available from http://beast.bio.ed.ac.uk/, 2006
48. F. Abascal, R. Zardoya and D. Posada, ProtTest: selection of best-fit models of protein evolution, *Bioinformatics*, 2005, **21**, 2104-2105.
49. G. A. O'Toole and R. Kolter, Initiation of biofilm formation in Pseudomonas fluorescens WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis, *Mol Microbiol*, 1998, **28**, 449-461.
50. F. W. Studier, Protein production by auto-induction in high density shaking cultures, *Protein Expr Purif*, 2005, **41**, 207-234.
51. M. T. Alexandre, J. C. Arents, R. van Grondelle, K. J. Hellingwerf and J. T. Kennis, A base-catalyzed mechanism for dark state recovery in the Avena sativa phototropin-1 LOV2 domain, *Biochemistry*, 2007, **46**, 3129-3137.
52. J. M. Christie, S. B. Corchnoy, T. E. Swartz, M. Hokenson, I. S. Han, W. R. Briggs and R. A. Bogomolni, Steric Interactions Stabilize the Signaling State of the LOV2 Domain of Phototropin 1, *Biochemistry*, 2007.
53. J. L. Hobman, J. Wilkie and N. L. Brown, A design for life: prokaryotic metal-binding MerR family regulators, *Biometals*, 2005, **18**, 429-436.
54. M. A. Schell, Molecular biology of the LysR family of transcriptional regulators, *Annu Rev Microbiol*, 1993, **47**, 597-626.
55. H. Takano, D. Asker, T. Beppu and K. Ueda, Genetic control for light-induced carotenoid production in non-phototrophic bacteria, *J Ind Microbiol Biotechnol*, 2006, **33**, 88-93.
56. K. W. Bayles, The biological role of death and lysis in biofilm development, *Nat Rev Microbiol*, 2007, **5**, 721-726.
57. K. C. Rice, E. E. Mann, J. L. Endres, E. C. Weiss, J. E. Cassat, M. S. Smeltzer and K. W. Bayles, The cidA murein hydrolase regulator contributes to DNA release and biofilm development in Staphylococcus aureus, *Proc Natl Acad Sci U S A*, 2007, **104**, 8113-8118.
58. P. Visca, F. Imperi and I. L. Lamont, Pyoverdine siderophores: from biogenesis to biosignificance, *Trends Microbiol*, 2007, **15**, 22-30.
59. K. Miyamoto, K. Nakahigashi, K. Nishimura and H. Inokuchi, Isolation and characterization of visible light-sensitive mutants of Escherichia coli K12, *J Mol Biol*, 1991, **219**, 393-398.
60. G. C. Ferreira, R. Franco, S. G. Lloyd, I. Moura, J. J. Moura and B. H. Huynh, Structure and function of ferrochelatase, *J Bioenerg Biomembr*, 1995, **27**, 221-229.
61. Z. Qi and M. R. O'Brian, Interaction between the bacterial iron response regulator and ferrochelatase mediates genetic control of heme biosynthesis, *Mol Cell*, 2002, **9**, 155-162.
62. C. Baysse, S. Matthijs, M. Schobert, G. Layer, D. Jahn and P. Cornelis, Co-ordination of iron acquisition, iron porphyrin chelation and iron-protoporphyrin export via the cytochrome c biogenesis protein CcmC in Pseudomonas fluorescens, *Microbiology*, 2003, **149**, 3543-3552.

4 Application of bacterial LOV proteins and future persepective

4.1 Reporter proteins for *in vivo* fluorescence without oxygen

Drepper T., Eggert T., Circolone F., Heck A., Krauss U., Guterl J.K., Wendorff M., Losi A., Gärtner W. and Jaeger K.E.
Nat. Biotechnol. (2007), 25: 443 - 445.

4.2 *insilico.mutagenesis*: a primer selection tool designed for sequence scanning applications used in directed evolution experiments

Krauss U. and Eggert T.
Biotechniques (2005) 39: 679 - 682.

Reporter proteins for *in vivo* fluorescence without oxygen

Thomas Drepper[1,5], Thorsten Eggert[1,4,5], Franco Circolone[1], Achim Heck[1], Ulrich Krauß[1], Jan-Karl Guterl[1], Marion Wendorff[1], Aba Losi[2], Wolfgang Gärtner[3] & Karl-Erich Jaeger[1]

Fluorescence reporter proteins like the green fluorescent protein (GFP) are applied as non-invasive molecular tools for labeling living specimen. The usage of these fluorescence reporters is generally restricted to aerobic systems, as the formation of their chromophores strictly depends on oxygen[1,2]. Here we describe the creation of novel anaerobically fluorescing proteins derived from two bacterial FMN-binding photoreceptors, namely *Bacillus subtilis* YtvA[3] and *Pseudomonas putida* SB2[4] which belong to the family of LOV (light oxygen voltage) proteins[5]. We have used the facultative anaerobic bacterium *Rhodobacter capsulatus* to demonstrate that these FMN-based fluorescent proteins (FbFP) can be applied as in vivo fluorescence reporters: While the yellow fluorescent protein (YFP) was detectable only in aerobically grown *R. capsulatus* cells, fluorescence of the novel FbFPs could clearly be detected under both aerobic and anaerobic conditions. Thus, FbFPs open up a broad spectrum of in vivo applications allowing to obtain novel insights into anaerobic life.

Reprinted by permission from Macmillan Publishers Ltd: Nature Biotechnology, **25**: 443-5, Copyright (2007)
Note: Supplementary information is available on the Nature Biotechnology website
Note: This manuscript represents the author's version of the peer-reviewed manuscript and is thus not identical to the copyedited version of the article available on the Nature Biotechnology website.

Today, genetically encoded fluorescent probes like GFP from the jellyfish *Aequorea victoria* have become invaluable tools for in vivo real-time imaging of cells and tissues. The development of various sophisticated GFP variants as well as the isolation of other GFP related proteins from marine organisms[6] has made the GFP family of fluorescent proteins to one of the cornerstones for in vivo labeling applications. GFP-like proteins can be used to analyze the expression, localization, movement, and interaction of proteins[2,7-10], thereby facilitating the analysis of complex cellular processes as well as biotechnological and biomedical applications[11,12]. Although these fluorescent proteins obviously possess many advantages for in vivo as well as in vitro fluorescence labeling, they also share at least one major drawback: all members of the GFP family strictly depend on molecular oxygen as a cofactor for the synthesis of their respective chromophores[1,2]. This limitation has prompted us to construct a set of novel FMN-based fluorescent proteins (FbFP) which allow in vivo labeling and detection in the presence and absence of oxygen. For that purpose the bacterial blue-light photoreceptor YtvA from *B. subtilis*[3] and SB2 from *P. putida*[4] were used to engineer reporter proteins fluorescing in the absence of oxygen. YtvA is a 261 amino acid protein consisting of two functional domains, the photoactive N-terminal LOV-domain (amino acid residue 25-126) and the C-terminal STAS domain (sulfate transporter / anti-sigmafactor antagonist, amino acid residue 147-258) carrying a nucleoside triphosphate (NTP) binding motif[13]. In *B. subtilis*, YtvA is involved in a complex signaling pathway that controls the general environmental stress response and acts as a positive regulator of the alternative transcription factor σ^B in response to blue light[14-16]. In contrast to YtvA, the putative photoreceptor SB2 from *P. putida* consists of 151 amino acids and forms a single LOV domain whereas a fused effector domain is missing[4]. Presently, the physiological function of the *P. putida* photoreceptor is unknown.

Conformational analysis revealed that the SB2 wild type protein as well as a truncated YtvA derivative YLOV (solely consisting of the LOV domain) forms stable homodimers, whereas full-length YtvA occurs as a monomer in solution[4,17]. The LOV domain which belongs to the structurally conserved PAS (PerArntSim) superfamily[5] was first characterized in plant blue-light receptors (phototropins, Phot) and binds flavin mononucleotide (FMN) chromophores noncovalently (**Supplementary Fig. 1f**).

[1]Institute of Molecular Enzyme Technology, Heinrich-Heine-University Duesseldorf, Research Centre Juelich, D-52426 Juelich, Germany
[2]Department of Physics, University of Parma, I-43100 Parma, Italy, [3]Max-Planck-Institute for Bioinorganic Chemistry, D-45470 Mülheim an der Ruhr; Germany, [4]EVOcatal GmbH, Merowingerplatz 1a D-40225 Duesseldorf, Germany
[5]These authors contributed equally to the work

When irradiated with blue light (450 nm) plant phototropins as well as YtvA and SB2 basically undergo the same photocycle involving the reversible formation of an FMN-cysteine C(4a)-thiol adduct and exhibit a weak intrinsic autofluorescence which occurs during the photocycle with a maximal emission wavelength of 495 nm[18,19,3,4].

Previous studies have demonstrated that the mutational replacement of the conserved FMN-binding cysteine residue with a non-polar alanine in the LOV2 domain of Phot1 from *Arabidopsis* and *Avena sativa* (oat) resulted in an increase of fluorescence emission[18,20]. The observed increase was due to a reduction of fluorescence quenching of the FMN chromophore as well as inhibition of the LOV2 photocycle.

In order to enhance the natural but weak autofluorescence of the bacterial photoreceptors we first substituted by site directed mutagenesis the conserved photoactive Cys62 of YtvA and Cys53 of SB2 by an alanine henceforth designated as *B. subtilis* (Bs) and *P. putida* (Pp) FbFP. The nucleotide and amino acid sequences of BsFbFP and PpFbFP are shown in **Supplementary Fig. 1e**. The resulting LOV derivatives were expressed in *E. coli* and *R. capsulatus*, respectively, using the newly constructed broad host range expression plasmid pRcExp1II which contains an inducible T7-promoter (Drepper *et al.*, manuscript in preparation). Comparative *in vivo* fluorescence measurements were first carried out in *E. coli* liquid batch cultures (**Fig. 1)**. Upon exposure to blue light (365 nm) *E. coli* cells expressing the engineered YtvA protein BsFbFP exhibited a significant increase of *in vivo* fluorescence as compared to cells expressing the respective wild type protein YtvA (**Fig. 1a**). Quantification of BsFbFP mediated *in vivo* fluorescence intensity revealed a 10-fold increase (**Fig. 1b**) whereas the expression level and the stability of the YtvA variant were not affected (**Fig. 1c,** lanes 2 and 3). In addition, the *in vivo* fluorescence intensity of the mutated *P. putida* LOV protein PpFbFP expressed in *E. coli* was similar to the intensity observed with BsFbFP (**Fig. 1a**,**b**).

In an attempt to further enhance the BsFbFP mediated fluorescence in *E. coli*, the DNA sequence of a truncated gene solely encoding the photoactive LOV domain (i.e. the first 137 N-terminal amino acid residue of BsFbFP) was adjusted to the *E. coli* codon bias (sequences are shown in **Supplementary Fig. 1e** online). Remarkably, the expression of the resulting protein named EcFbFP dramatically improved the fluorescence of *E. coli* cells (**Fig. 1a**) and the fluorescence intensity in batch cultures was increased about 2.5-fold in comparison to variant BsFbFP and about 25-fold when compared to YtvA wt, respectively (**Fig. 1b**). Surprisingly, the change of the codon bias did not result in an increased amount of EcFbFP (**Fig. 1c**, lane 4).

To analyze the fluorescence properties of the novel reporter proteins in more detail, PpFbFP and EcFbFP were heterologously expressed in *E. coli* BL21 (DE3) as C-terminal His-tag fusion proteins. After purification, the protein-to-chromophore ratio was determined by UV/VIS (272/447 nm) spectroscopy for EcFbFP. As described for YtvA wild-type protein[21], the UV/VIS absorbance ratio of EcFbFP was 4.2 corresponding to a 1:1 ratio of protein to chromophore, although the expression level of EcFbFP was very high with a yield of purified protein of 35.5 mg per liter of *E. coli* culture. Thus, the supply of endogenous FMN is not limiting the use of FbFPs as *in vivo* fluorescence reporter proteins. The light absorption and fluorescence excitation spectra of the recombinant LOV proteins correlated well with the spectra of the corresponding wild type proteins (**Fig. 1d,** [3,4]). Both proteins showed maximal light absorption at a wavelength of 449 nm and a maximal emission at 495 nm upon excitation with blue-light (450 nm). For reasons discussed previously by Losi and coworkers[3], the molar extinction coefficient of free FMN which is 12,500 ± 500 M^{-1} cm^{-1} at 450 nm was used instead of that of the purified LOV-proteins. Determination of FbFP fluorescence quantum yields has been carried out by using FMN (Q_F = 0.26) as a reference standard revealing a quantum yield of 0.17 for PpFbFP and 0.39 for EcFbFP, respectively. Therefore, the relative fluorescence brightness of the FbFPs is in the same range as determined for the blue and cyan fluorescent variants of GFP (BFP: extinction coefficient = 31,000 and Q_F = 0.25; CFP: extinction coefficient = 26,000 and Q_F = 0.40)[22]. Finally, an initial analysis of PpFbFP and EcFbFP photobleaching properties showed that both fluorescent proteins exhibit a high photostability, with the photobleaching obeying a double exponential decay with τ_1=909 s (relative Amplitude (A) A1 = 49%), τ_2=1835 s (A2 = 51%) for PpFbFP and τ_1=67 s (A1 = 98%), τ_2=15135 s (A2 = 2%) for EcFbFP. Accordingly, these exponential time constants can be converted to the respective quantum yields Q1=5.09E-8 and Q2=2.52E-8 for $\tau_1$and τ_2 of PpFbFP and Q1=6.91E-7 and Q2=3.06E-9 for τ_1 and τ_2 of EcFbFP. Virtually all LOV proteins of prokaryotic and eukaryotic origin contain FMN as their chromophore[5]. Since the synthesis of FMN is independent of oxygen, we further tested if the novel LOV-based fluorescent proteins can be used as genetically encoded fluorescent probes under strictly anaerobic conditions. Anaerobically fluorescing proteins would be highly useful reporters in numerous free-living and biofilm-forming microbes that grow under oxygen limited or strictly anaerobic conditions in natural and artificial environments, e.g. in anaerobic digesters, biodegradation niches or industrial reactors as well as in hypoxic regions and tissues of higher organisms such as a variety of human solid tumors[23].

Our laboratory is concerned with sustainable biotechnological applications which are based on the facultative anaerobic, phototrophic bacterium *R. capsulatus*[24,25]. Therefore, we used this organism as an appropriate test system to analyze *in vivo* fluorescence mediated by the engineered LOV proteins in relation to oxygen availability. Thus, LOV expression plasmids pRcExp1II-EcFbFP and pRcExp1II-PpFbFP were transferred into the expression strain *R. capsulatus* B10S-T7 (Drepper *et al.*, manuscript in preparation). For comparison, the gene encoding the GFP derivative YFP was cloned into the same broad host range expression vector resulting in plasmid pRcExp1II-YFP which was parallel transferred to *R. capsulatus* B10S-T7.

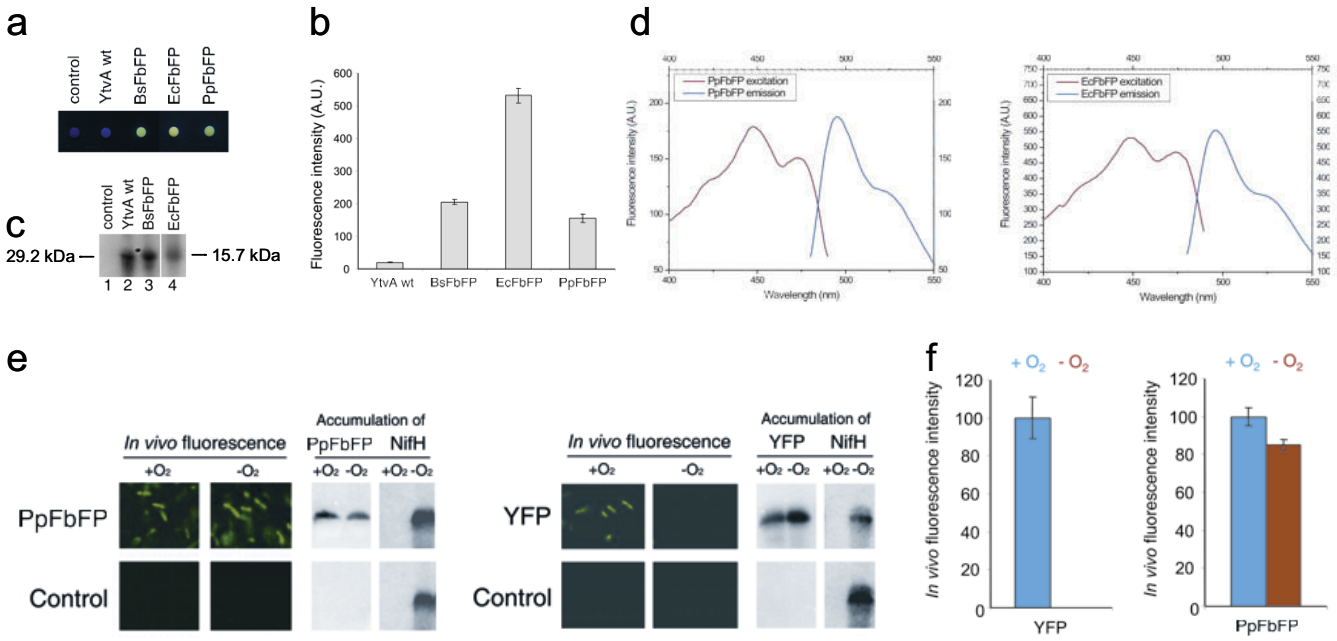


Figure 1 Novel fluorescence reporter proteins created from LOV-proteins. The experiments were conducted using *E. coli* BL21 (DE3) cells carrying a T7 polymerase-based broad host range expression plasmid (pRcExp1II) encoding the indicated YtvA and SB2 variants. (**a**) *In vivo* fluorescence of different engineered LOV proteins in *E. coli*. Cells expressing the indicated LOV protein variants were grown until stationary phase in LB media containing 1.0 mM IPTG. *In vivo* fluorescence was visualized by illuminating drops of the respective cell suspensions under blue light (365 nm). (**b**) Quantification of FbFP mediated autofluorescence. *E. coli* expression strains were grown in LB medium and expression of FbFPs was specifically induced by adding 0.4 mM IPTG. Fluorescence intensity of cell extracts was recorded at 495 nm upon excitation with blue light (450 nm). Values are means of triplicate measurements with background fluorescence from *E. coli* BL21 (pRcExp1II) subtracted and normalized against cell density. A.U.: Arbitrary Units. (**c**) FbFP accumulation in *E. coli*. Accumulation of fluorescent proteins was analyzed by using the same cell extracts as in (**b**). For the immuno detection of LOV derivatives polyclonal antisera raised against wild type YtvA and SB2 were used. **(d)** Fluorescence emission spectra (in blue) of purified PpFbFP and EcFbFP recorded at an excitation wavelength of 450 nm. Fluorescence excitation spectra (shown in red) of the same samples were recorded at an emission wavelength of 495 nm. All measurements were conducted in 10mM Na-phosphate buffer pH 8.0, 10mM NaCl at 20°C with protein samples diluted to approximately 400µg/ml. A.U.: Arbitrary Units. Comparison of YFP and PpFbFP mediated *in vivo* fluorescence in the absence and presence of oxygen. *R. capsulatus* B10S-T7 cells carrying expression plasmid pRcExp1II encoding either YFP (**e**) or the C53A variant of SB2 (**e**) were grown under aerobic ($+O_2$) as well as anaerobic ($-O_2$) conditions and *in vivo* fluorescence was analyzed microscopically using a Zeiss LSM 510 confocal laser scanning microscope. Cells were excited with laser light at defined wavelengths of 488 nm (YFP) and 458 nm (PpFbFP). Fluorescence emission was detected in the ranges of 505-550 nm (YFP) and 475-525 nm (PpFbFP), respectively. Control experiments with *R. capsulatus* cells without a fluorescence reporter gene were accomplished in the same way. Accumulation of YFP and PpFbFP in *R. capsulatus* was analyzed by immuno detection using specific antisera raised against YFP and SB2, respectively. (**f**) For whole-cell liquid culture fluorescence measurements the same cell cultures as in (**e**) were used. Fluorescence emission of living cells was recorded at 529nm (YFP) and 495 nm (PpFbFP). Excitation wavelengths were 488 nm for YFP and 450 nm for PpFbFP. Fluorescence intensities represent averages of three independent measurements with background fluorescence from *R. capsulatus* B10S-T7 (pRcExp1II) subtracted and normalized against cell density. Fluorescence intensities of YFP and PpFbFP which could be achieved in the presence of oxygen were used as a reference (100%) for the respective anaerobic fluorescence.

The negative control consisted of *R. capsulatus* B10S-T7 carrying the expression vector without insert. The *R. capsulatus* expression strains were grown in RCV minimal medium under aerobic as well as anaerobic conditions. After induction of the reporter gene expression the effect of oxygen on fluorescence and accumulation of YFP and FbFPs was analyzed by using confocal laser scanning microscopy and by immuno detection (**Fig. 1e**). YFP mediated fluorescence could only be monitored in living *R. capsulatus* cells in the presence of oxygen as the synthesis of the chromophore is fully dependent on the availability of molecular oxygen[1]. Nevertheless, equal amounts of YFP protein were synthesized under aerobic ($+O_2$) and anaerobic ($-O_2$) growth conditions (**Fig 1e**). In contrast, PpFbFP mediated fluorescence emission signals were clearly detectable in *R. capsulatus* cells grown under both anaerobic and aerobic conditions (**Fig. 1e**). The levels of EcFbFP mediated fluorescence and protein accumulation were very low in *R. capsulatus* indicating an inefficient gene expression and/or a rapid degradation of the recombinant protein in the phototrophic bacterium (data not shown). The O_2-independent fluorescence of variant PpFbFP was finally confirmed in the respective *R. capsulatus* expression cultures by demonstrating the accumulation of dinitrogenase reductase (NifH) which is known to be synthesized in *R. capsulatus* only under strictly anaerobic growth conditions[26,27]. **Figure 1e** clearly shows that the NifH protein was synthesized exclusively under anaerobic conditions thereby corroborating the absence of O_2 under the anaerobic assay conditions.The YFP and PpFbFP mediated fluorescence was quantified in the presence and absence of O_2 in living cells of the *R. capsulatus* expression strains by measuring fluorescence in whole-cell liquid cultures (**Fig. 1f**). Aerobic and

anaerobic *R. capsulatus* cultures were obtained by growing the respective YFP and PpFbFP expression strains until late logarithmic phase. To avoid any aerobic recovery of YFP fluorescence during cell harvesting[28] fluorescence quantification was carried out within one minute. Under these experimental conditions, YFP dependent fluorescence was completely abolished in the absence of oxygen (**Fig. 1f**, $-O_2$) whereas bright *in vivo* fluorescence was observed in the presence of oxygen (**Fig. 2c**, $+O_2$). In contrast, the absence of oxygen did not lead to a significant decrease of *in vivo* fluorescence intensity in *R. capsulatus* cells expressing the LOV variant PpFbFP (**Fig. 1f**).In this communication, we describe the construction of novel FMN-based fluorescing proteins (FbFPs) which originate from bacterial blue-light photoreceptors. In contrast to coelenterate fluorescent proteins and LuxY which is the only known natural occurring FMN-binding fluorescent protein from the bioluminescent marine bacterium *Vibrio fischeri*[29,30], the engineered FbFPs do not lose fluorescence activity under oxygen deprivation. Thus, these FbFPs can solve longstanding problems of protein-probed fluorescent imaging in anaerobic systems by providing experimental tools to expand the existing *in vivo* fluorescence imaging techniques. Furthermore, the study of complex anaerobic processes by using FbFPs as fluorescence reporter proteins holds a tremendous potential for a variety of different biotechnological and biomedical applications including: (i) high throughput screening of anaerobic bacteria by using fluorescence-activated flow cytometry and cell sorting (FACS), (ii) whole-cell biosensor based toxicity detection and microbial anaerobic detoxification, and (iii) biomedical diagnostics and development of novel cancer therapies by using strictly anaerobic bacteria as anti tumor agents[31].

ACKNOWLEDGMENTS

The authors would like to thank Jürgen Hubbuch and Thomas Rosenbaum, Institute of Biotechnology 2, Forschungszentrum Jülich, Germany for their assistance with confocal laser scanning microscopy. This work was partly supported by grants from the Deutsche Forschungsgemeinschaft (DFG) within the Forschergruppe FOR526 „Sensory blue light photoreceptors“

1. Tsien, R.Y. The Green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509-544 (1998).
2. Zimmer, M. Green flourescent protein (GFP): application, structure, and related photophysical behavior. *Chem. Rev.* **102**, 759-781 (2002).
3. Losi, A., Polverini, E., Quest, B. & Gärtner, W. First evidence for phototropin-related blue-light receptors in prokaryotes. *Biophys. J.* **82**, 2627-2634 (2002).
4. Krauss, U., Losi, A., Gärtner, W., Jaeger, K.E. & Eggert, T. Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct. *Phys. Chem. Chem. Phys.* **21**, 2804-2811 (2005).
5. Losi, A. The bacterial counterparts of pant phototropins. *Photochem. Photobiol. Sci.* **3**, 566-574 (2004).
6. Matz, M.V., Lukyanov, K.A. & Lukyanov, S.A. Family of the green fluorescent protein: journey to the end of the rainbow. *Bioessays* **24**, 953-959 (2002).
7. Chudakov, D.M., Lukyanov, S. & Lukyanov, K.A. Fluorescent proteins as a toolkit for *in vivo* imaging. *Trends Biotechnol.* **23**, 605-613 (2005).
8. Giepmans, B.N.G., Adams, S.R., Ellisman, M.H. & Tsien, R.Y. The fluorescent toolbox for assessing protein localisation and function. *Science* **312**, 217-224 (2006).
9. O'Rourke, N.A., Meyer, T. & Chandy, G. Protein localization studies in the age of 'omics'. *Curr. Opin. Chem. Biol.* **9**, 82-87 (2005).
10. Valdivia, R.H., Cormack, B.P. & Falkow, S. The uses of green fluorescent protein in prokaryotes. *Methods Biochem. Anal*. **47**, 163-178 (2006).
11. March, J.C., Rao, G. & Bentley, W.E. Biotechnological applications of green fluorescent protein. *Appl. Microbiol. Biotechnol.* **62**, 303-315 (2003).
12. Hassan, M. & Klaunberg, B.A. Biomedical applications of fluorescence imaging *in vivo*. *Comp. Med.* **54**, 635-644 (2004).
13. Aravind, L. & Koonin, E.V. The STAS domain a link between anion transporters and antisigma-factor antagonists. *Curr. Biol.* **10**, R53-R55 (2000).
14. Akbar, S. *et al.* New family of regulators in the environmental signalling pathway which activates the general stress transcription factor σ^B of *Bacillus subtilis*. *J. Bacteriol.* **183**, 1329-1338 (2001).
15. Gaidenko, T.A., Kim, T.J., Weigel, A.L., Brody, M.S. & Price, C.W. The blue-light receptor YtvA acts in the environmental stress signaling pathway of *Bacillus subtilis*. *J Bacteriol.* **188**, 6387-6395 (2006).
16. Avila-Perez, M., Hellingwerf, K.J. & Kort, R. Blue light activates the σ^B-dependent stress response of *Bacillus subtilis* via YtvA. *J. Bacteriol.* **188**, 6411-6414 (2006).
17. Buttani, V. *et al.* Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions. *Photochem. Photobiol. Sci*. **6**, 41-49 (2007).
18. Schwartz, T.E. *et al.* The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J. Biol. Chem.* **276**, 36493-36500 (2001).
19. Kasahara, M. *et al.* Photochemical properties of the flavin mononucleotide-binding domains of the phototropins from *Arabidopsis*, rice, and *Chlamydomonas reinhardtii*. *Plant Physiol.* **129**, 762-773 (2002).
20. Christie, J.M., Swartz, T.E., Bogomolni, A. & Winslow, R.B. Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function. *Plant J.* **32**, 205-219 (2002).
21. Losi, A., Ghiraldelli, E., Jansen, S. & Gärtner, W. Mutational effects on protein structural changes and interdomain interactions in the blue-light sensing LOV protein YtvA. *Photochem. Photobiol.* **81**, 1145-1152 (2005).
22. Patterson, G., Day, R.N. & Piston, D. Fluorescent protein spectra. *J. Cell Sci*. **114**, 837-838 (2001).

23. Brown, J.M. Exploiting the hypoxic cancer cell: mechanisms and therapeutic strategies. *Mol. Med. Today* **6**, 157-162 (2000).
24. Drepper, T., Arvani, S., Rosenau, F., Wilhelm, S. & Jaeger, K.-E. High-level transcription of large gene regions: a novel T_7 RNA-polymerase-based system for expression of functional hydrogenases in the phototrophic bacterium *Rhodobacter capsulatus*. *Biochem. Soc. Trans.* **33**, 56-58 (2005).
25. Drepper, T. *et al.* Novel biocatalysts for white biotechnology. *Biotechnol. J*. **1**, 777-786 (2006).
26. Drepper, T. *et al.* Regulation of synthesis and activity of molybdenum nitrogenase and the alternative nitrogenase in *Rhodobacter capsulatus*. In: *Regulatory Networks in Prokaryotes.* Horizon Scientific Press, Norfolk, U.K. 73-77 (2003).
27. Masepohl, B., Drepper, T. & Klipp, W. Nitrogen fixation in the phototrophic purple bacterium *Rhodobacter capsulatus*. In: *Genetics and regulation of nitrogen fixation in free-living bacteria.* Kluwer Academic Publischer, Dordrecht, Boston, London, 141-173 (2004).
28. Zhang, C., Xing, X.-H. & Lou, K. Rapid detection of a gfp-marked *Enterobacter aerogenes* under anaerobic conditions by aerobic fluorescence recovery. *FEMS Microbiol. Lett.* **249**, 211-218 (2005).
29. Baldwin, T.O., Treat, M.L. & Daubner, S.C. Cloning and expression of the *luxY* gene from *Vibrio fischeri* strain Y-1 in *Escherichia coli* and complete amino acid sequence of the yellow fluorescent protein. *Biochemistry* **29**, 5509-5515 (1990).
30. Karatani, H., Yoshizawa, S. & Hirayama, S. Oxygen triggering reversible modulation of *Vibrio fischeri* strain Y1 bioluminescence *in vivo*. *Photochem. Photobiol*. **79**, 120-125 (2004).
31. Bettegowda, C. *et al*. The genome and transcriptomes of the anti-tumor agent *Clostridium novyi-NT*. Nat. Biotechnol. Published online: 19 November 2006; doi:10.1038/nbt1256
32. Kovach, M.E., Phillips, R.W., Elzer, P.H., Roop, R.M. 2nd & Peterson, K.M. pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* **16**, 800-802 (1994).
33. Arnold, W. & Pühler, A. A family of high-copy-number plasmid vectors with single end-label sites for rapid nucleotide sequencing. *Gene* **70**, 171-179 (1988).
34. Klipp, W., Masepohl, B. & Pühler, A. Identification and mapping of nitrogen fixation genes of *Rhodobacter capsulatus*: duplication of a *nifA-nifB* region. *J. Bacteriol*. **170**, 693-699 (1988).
35. Masepohl, B., Klipp, W. & Pühler, A. Genetic characterization and sequence analysis of the duplicated *nifA*/*nifB* gene region of *Rhodobacter capsulatus*. *Mol. Gen. Genet.* **212**, 27-37 (1988).

insilico.mutagenesis: a primer selection tool designed for sequence scanning applications used in directed evolution experiments

Ulrich Krauss and Thorsten Eggert
Heinrich-Heine-Universität Düsseldorf, Jülich, Germany

BioTechniques 39:679-682 (November 2005)
doi 10.2144/000112013

*Several primer prediction programs have been developed for a variety of applications. However, none of these tools allows the prediction of a large set of primers for whole gene site-directed mutagenesis experiments using the megaprimer method. We report a novel primer prediction tool (*insilico.mutagenesis*), accessible at www.insilico.uni-duesseldorf.de, developed for the application to high-throughput mutagenesis used in directed evolution or structure-function dependency projects, which involve the subsequent mutagenesis of a large number of amino acid positions (e.g., in whole gene saturation or gene scanning mutagenesis experiments). Furthermore, the program is suitable for all site-directed (saturation) mutagenesis approaches, such as saturation mutagenesis of promoter sequences and other types of untranslated intergenic regions. In anticipation of downstream cloning steps, the primer design tool also includes a restriction site control feature alerting the user if unwanted restriction sites have been introduced within the mutagenesis primer. The use of our tool promises to speed up the process of site-directed mutagenesis, as it instantly allows predicting a large set of primers.*

INTRODUCTION

The design of PCR primers (i.e., single-stranded DNA oligonucleotides complementary to a target sequence) has become an essential procedure in molecular biology for a variety of applications (e.g., gene amplification, sequencing, and site-directed mutagenesis). Thus, numerous primer prediction and analysis programs have been developed (1–4). However, none of the programs currently available facilitate the design of primers for whole gene site saturation and sequence-scanning mutagenesis experiments using the megaprimer PCR method (5). The manual design of oligonucleotide PCR primers for these approaches is a laborious task, due to the large set of mutagenesis primers to be designed. In order to reduce the working time and to standardize the design procedure, we have developed a web-based primer prediction program termed *insilico.mutagenesis*. The program takes nucleotide sequences of target regions (open reading frames or intergenic regions like promoter sequences) as input and predicts mutagenesis primers that can be directly used in a megaprimer PCR-based mutagenesis approach. The *insilico.mutagenesis* tool is entirely written in Perl (6), uses MySQL tables for easy data storage, and possesses an HTML-based user interface. The tool can be accessed via the *insilico* web site (www.insilico.uni-duesseldorf.de).

Directed evolution has been proven to be a successful strategy to improve enzyme properties such as specific activities, substrate specificities, thermostabilities, or enantioselectivities (7–10). In most cases, single base mutations are introduced by means of error-prone PCR (epPCR) in a random manner. However, epPCR only results in a limited number of amino acid exchanges; therefore, only a small part of the total sequence space is accessible to mutagenesis. As a consequence, alternative techniques must be applied to generate a first generation library of high diversity (11). Complete saturation mutagenesis, also referred to as Gene Site Saturation Mutagenesis™ (GSSM™), is a novel technology for rapid in vitro evolution of proteins that can be used to circumvent this problem (12–14). Here, all possible base triplets are introduced at a given codon position, thereby resulting in the formation of a library containing all 20 amino acid exchanges at the target position. This is achieved at the genetic level by using degenerate mutagenesis primers. Subsequent use of in vitro PCR amplification generates a library of genes possessing all codon variations required for complete saturation of the original gene. DeSantis and coworkers applied this technique to generate a highly enantioselective nitrilase (13). Furthermore, the technique of complete saturation mutagenesis has been used in our institute to generate a variant *Bacillus subtilis* lipase A (BLSA) showing improved enantioselectivity toward different model substrates (14,15). Sequence-scanning mutagenesis techniques, like alanine- or tryptophan-scanning mutagenesis, can be applied to investigate the functional role of specific amino acid residues with respect to catalytic mechanism, substrate binding, or signal transduction (16,17).

Both complete saturation mutagenesis and scanning-mutagenesis techniques require the sequential saturation/substitution of numerous amino acid residues, depending on the size of the target protein or the region to be investigated. For the complete saturation of a regular protein consisting of 300 amino acids, 300 single codon exchanges (i.e., 300 megaprimer PCRs) must be performed. One step in such a challenging approach that is easily amenable to automation without the necessity of expensive robotic equipment, is the primer design using a personal computer. Therefore, we developed the program *insilico.mutagenesis* to automate the prediction of oligonucleotides that can be used directly in a megaprimer PCR approach.

REQUIRED INPUT

A schematic overview of the data-processing by *insilico.mutagenesis* is given in Figure 1. First, the program

requires the input of a target nucleotide sequence, including flanking vector sequences (plain sequence), to which we will refer as vectorA-template. It is not necessary to include the complete vector-sequence; about 40 bp up- and downstream of the gene of interest is enough to enable primer design for whole gene saturation or scanning mutagenesis. Second, a unique sequence identifier (sequence name) must be provided for data processing purposes. Third, the mutagenesis codon must be selected from a pull-down menu, taking into account the codon-usage of the desired expression host. Next, the program requires the input of the start and stop position of the target gene (or intergenic region) within the overall sequence. Also, the region that should be mutated must be specified, and because, in practice, a too long megaprimer might be inefficiently elongated by the polymerase in the second round of PCR—probably due to the formation of secondary structures—it has been proven practical to design mutagenesis primers in a way that the megaprimer does not exceed a certain length (5). Therefore, the input of a so-called oligo-switch position is necessary, as explained in more detail later.

PROGRAM ALGORITHM AND DATA PROCESSING

As one example, we designed all mutagenesis primers in the complete saturation mutagenesis of a *B. subtilis* lipase (14). The gene was 543 bp in length, consequently having 181 coding triplets and a TAA stop codon. The first 90 codon-exchanges (corresponding to 270 bp) are achieved by the design of reverse mutagenesis primers (as shown in Figure 2B), which are used together with a vectorA-specific forward primer. Accordingly, the last 91 codon-exchanges are introduced using a forward mutagenesis primer together with a vectorA-specific reverse primer. As a consequence, the amplified megaprimers do not exceed 273 bp in size. Therefore, these DNA fragments are well suited for the second PCR of megaprimer mutagenesis (5,18). The position at which the "switch" from

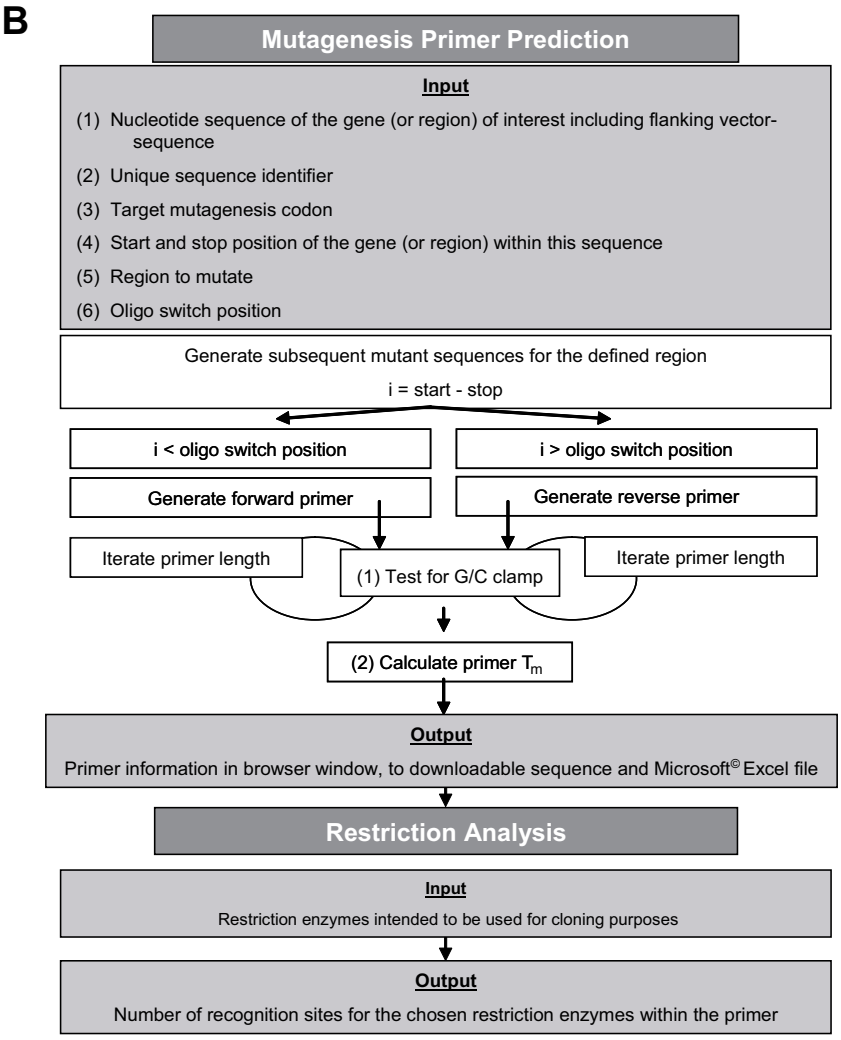


Figure 1. User interface and flowchart of data processing of the *insilico.mutagenesis* primer prediction tool. (A) Screenshot of the tool's data input user interface. (B) Flowchart of the tool's data processing and data generation algorithm. T_m, melting temperature.

a reverse to a forward mutagenesis primer occurs is referred to as the oligo-switch position. Usually, as in our lipase example, the position halfway along the gene of interest is used.

Using this input (Figures 1 and 2) the program generates mutant nucleotide sequences with subsequent single codon-exchanges within the region the user has defined. For each of the mutant sequences, a mutagenesis primer is predicted either as a forward or reverse primer, depending on the position of the desired mutation with respect to the oligo-switch position. The generated mutagenesis primers are checked for the ability to form the so-called GC clamps at the 3′ end, since Watson-Crick bonds between G and C will facilitate the initiation of complementary strand formation by the polymerase at the 3′ end of the hybridized primer (19). If no GC clamps can be formed due to the lack of G or C bases at the 3′ end of the primer, the program extends the oligonucleotide until a G or C is found at its 3′ end. The maximum length of the primer is set at 40 bp. Finally, the mutagenesis primer data are stored in a MySQL database and displayed in form of an HTML table (Figure 1). In addition, the program calculates the melting temperature of every single primer based on the equation of Breslauer et al. (20) and the nearest neighbor thermodynamic parameter set as described by Allawi and SantaLucia (21). Furthermore, the predicted primer sequences can be viewed as FASTA-formatted text output in the browser window or can be downloaded as a Microsoft® Excel® spreadsheet.

ADDITIONAL ANALYSES

The program enables the user to check each predicted mutagenesis primer with respect to additional restriction endonuclease recognition sites, which might interfere with the intended cloning strategy for the amplified full-length PCR product. Therefore, the program asks the user to supply the names of two restriction enzymes that will be used in the subsequent cloning steps. By using the BioPerl (22) module Bio::Restriction::Analysis, *insilico.mutagenesis* indicates the number of recognition sites of those enzymes within each oligonucleotide. In case the program has predicted a mutagenesis primer whose sequence interferes with the desired cloning strategy, the oligonucleotide can be redesigned easily.

CONCLUSIONS

In summary, we have presented a novel primer design tool (*insilico.mutagenesis*) specifically developed for high-throughput mutagenesis primer prediction, useful in complete saturation and whole gene scanning mutagenesis experiments. Thus, *insilico.mutagenesis* is designed to speed up the process of directed evolution or structure-function dependency projects. Furthermore, the primer design tool includes a restriction site control feature alerting the user in case of introducing unwanted restriction sites within the mutagenesis primer anticipating the cloning strategy.

ACKNOWLEDGMENTS

We thank Bernd Cappel (HHU Düsseldorf) for help installing the program in a server-based environment to make it accessible via the world wide web.

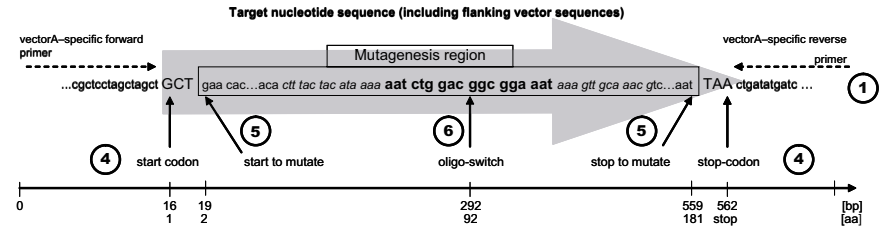
COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

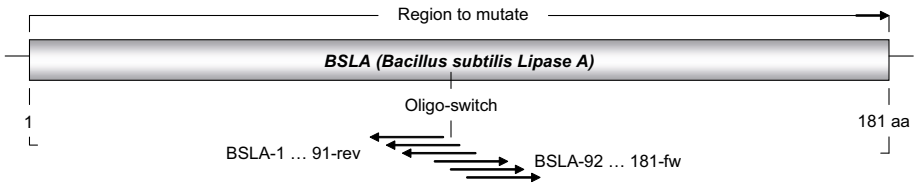
REFERENCES

1. **Rozen, S. and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. *132*:365-386.
2. **Lu, G., M. Hallet, S. Pollock, and D. Thomas.** 2003. DePie: designing primers

A



B



Oligonucleotide name	Sequence	Oligonucleotide orientation	Length (bp)	T_m (°C)
BLSA-89-rev	5′- atttccgccgtccagSNNttttatgtagtaaag -3′	reverse	33	deg.
BLSA-90-rev	5′- tttatttccgccgtcSNNatttttatgtagtaaag -3′	reverse	36	deg.
BLSA-91-rev	5′- aactttatttccgccSNNcagatttttatgtag -3′	reverse	34	deg.
BLSA-92-fw	5′- ataaaaaatctggacNNSggaaataaagttg -3′	forward	31	deg.
BLSA-93-fw	5′- aaaaatctggacggcNNSaataaagttgcaaac -3′	forward	33	deg.
BLSA-94-fw	5′- aatctggacggcggaNNSaaagttgcaaacg -3′	forward	31	deg.

Figure 2. Application of *insilico.mutagenesis* to design primers for site-directed mutagenesis. (A) Visualization of the *Bacillus subtilis* Lipase A (BSLA) nucleotide sequence, explaining the input requirements of the *insilico.mutagenesis* tool. Codons 89–94 of the BSLA gene sequence are printed in bold face. (B) Orientation of mutagenesis primers around the oligo-switch position in the BSLA gene and sequences of the oligonucleotides computed by the *insilico.mutagenesis* tool. deg.,degenerate; T_m, melting temperature.

for protein interaction experiments. Nucleic Acids Res. *31*:3755-3757.
3. **Turchin, A. and J.F. Lawler, Jr.** 1999. The Primer Generator: a program that facilitates the selection of oligonucleotides for site-directed mutagenesis. BioTechniques *26*:672-676.
4. **Canaves, J.M., A. Morse, and B. West.** 2004. PCR primer selection tool optimized for high-throughput proteomics and structural genomics. BioTechniques *36*:1040-1042.
5. **Barettino, D., M. Eigenbutz, R. Valcarcel, and H.G. Stunnernberg.** 1994. Improved method for PCR-mediated site-directed mutagenesis. Nucleic Acids Res. *22*:541-542.
6. **Moorhouse, M. and P. Barry.** 2004. Bioinformatics, Biocomputing and Perl. John Wiley & Sons, New York.
7. **Petrounia, I.P. and F. Arnold.** 2000. Designed evolution of enzymatic properties. Curr. Opin. Biotechnol. *11*:325-330.
8. **Brakmann, S.** 2001. Discovery of superior enzymes by directed molecular evolution. ChemBioChem. *2*:865-871.
9. **Reetz, M.T.** 2004. Controlling the enantioselectivity of enzymes by directed evolution: practical and theoretical ramifications. Proc. Natl. Acad. Sci. USA *101*:5716-5722.
10. **Jaeger, K.-E. and T. Eggert.** 2004. Enantioselective biocatalysis optimized by directed evolution. Curr. Opin. Biotechnol. *15*:305-313.
11. **Eggert, T., M.T. Reetz, and K.-E. Jaeger.** 2004. Directed evolution by random mutagenesis: a critical evaluation, p. 375-390. *In* A. Svendsen (Ed.), Enzyme Functionality: Design, Engineering, and Screening. Marcel Dekker, New York.
12. **Short, J.M., inventor.** Diversa Corporation, assignee. 2001. Saturation mutagenesis in directed evolution. U.S. Patent No. 6.171.820.
13. **DeSantis, G., K. Wong, B. Farwell, K. Chatman, Z. Zhu, G. Tomlinson, H. Huang, X. Tan, et al.** 2003. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). J. Am. Chem. Soc. *125*:11476-11477.
14. **Funke, S.A., A. Eipper, M.T. Reetz, N. Otte, W. Thiel, G. van Pouderoyen, B.W. Dijkstra, K.-E. Jaeger, and T. Eggert.** 2003. Directed evolution of an enantioselective *Bacillus subtilis* lipase. Biocatal. Biotransformation *21*:67-73.
15. **Funke, S.A., N. Otte, T. Eggert, M. Bocola, K.-E. Jaeger, and W. Thiel.** Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. Protein Eng. Des. Sel. (In press).
16. **Mahan, S.D., G.C. Ireton, B.L. Stoddard, and M.E. Black.** 2004. Alanine-scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference. Biochemistry *43*:8957-8964.
17. **Santiago, J., G.R. Guzman, K. Torruellas, L.V. Rojas, and J.A. Lasalde-Dominicci.** 2004. Tryptophan scanning mutagenesis in the TM3 domain of the *Torpedo califonica* acetylcholine receptor beta subunit reveals an alpha-helical structure. Biochemistry *43*:10064-10070.
18. **Ling, M.M. and B.H. Robinson.** 1997. Approaches to DNA mutagenesis: an overview. Anal. Biochem. *254*:157-178.
19. **Lowe, T.M., J. Sharefkin, S.Q. Yang, and C.W. Dieffenbach.** 1990. A computer program for selection of oligonucleotide primers for the polymerase chain reaction. Nucleic Acids Res. *18*:1757-1761.
20. **Breslauer, K.J., R. Frank, H. Blocker, and L.A. Marky.** 1986. Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. USA *83*:3746-3750.
21. **Allawi, H.T. and J. SantaLucia, Jr.** 1997. Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry *36*:10581-10594.
22. **Stajich, J.E., D. Block, K. Boulez, S.E. Brenner, S.A. Chervitz, C. Dagdigian, G. Fuellen, J.G. Gilbert, et al.** 2002. The bioperl toolkit: Perl modules for live sciences. Genome Res. *12*:1611-1618.

Received 19 April 2005; accepted 3 June 2005.

Address correspondence to Thorsten Eggert, Heinrich-Heine-Universität Düsseldorf, Institut für Molekulare Enzymtechnologie, Forschungszentrum Jülich, D-52426 Jülich, Germany. e-mail: t.eggert@fz-juelich.de

5 General Discussion

5.1 LOV is all around – the ancient history of a omnipresent photoreceptor protein

As discussed in Chapter 1.6 and Chapter 2.1 several recent ground-breaking studies unequivocally demonstrated a blue-light dependent functionality of different bacterial LOV domain-containing signaling systems and hence proved that blue-light is indeed an important environmental stimulus also in non-photosynthetic prokaryotes.
Years after the discovery of the LOV photosensor modules in plant phototropins, experimental evidence becomes available stating that bacterial systems similar to plant phototropins, namely LOV histidine kinases, can regulate blue-light dependent responses via a phot-like autophosphorylation system. This raises pressing questions regarding the evolutionary history of the LOV photosensor module in the pro- and eukaryotic domains of life. Therefore, a phylogenetic study that included 129 gene/protein sequences with significant similarity to the plant LOV domains of the phototropin system was conducted. The results of this study aiming to understand the evolutionary history of the LOV domain signaling module in the kingdoms domains of life, namely the Archaea, the Bacteria and the Eukaryotes, are presented in Chapter 2.1.

LOV is all around

The LOV signaling module is found widely distributed in all three kingdoms of life and is apparently only absent in the animals. A comprehensive BLAST analysis as well as independent genome mining in various fungal genome projects revealed that the LOV signaling module can be found in approximately 12% of bacterial and 7% of archaeal completely sequenced genomes. In the resulting phylogenetic analysis furthermore representatives of the three major eukaryotic LOV signaling systems

(phototropin, ZTL/ADO/FKF1 and fungal white-collar-1) were included in this analysis to gain insight into the evolutionary relationship between the prokaryotic and the eukaryotic LOV-systems.

The bacterial origin of the eukaryotic LOV signaling systems

Previous sequence and BLAST analyses had already indicated that the LOV signaling module should represent an evolutionary ancient light sensor that apparently retained its photochemical mechanism from archaeal taxa, to bacterial and eukaryotic genera throughout evolution. Although, biochemical/photochemical evidence is still missing that the archaeal LOV-like proteins possess the same characteristics as shown for several bacterial and eukaryotic LOV domains and full-length photosensor proteins. Nevertheless, the high degree of sequence conservation, especially of mechanistically crucial amino acids also in the archaeal LOV proteins, suggests the conservation of the LOV signaling paradigm between the three kingdoms of life. However, no phylogenetic study was available that investigated the evolutionary relationship among the pro- and eukaryotic world of LOV sensor systems. The analyses presented in Chapter 2.1 suggest that all known eukaryotic LOV photosensor families, namely, phototropins, ZTL/ADO/FKF1-LOVs and the fungal WC-1 photoreceptors have probably directly originated from cyanobacterial (in case of the ZTL-subfamily) or respectively α-proteobacterial lineages (in case of phototropin and WC-1) by endosymbiosis. Therefore, the respective endosymbiotic uptake event that led globally to the appearance of the organelles (mitochondria and chloroplast) in the eukaryotes apparently also marks the time of the appearance of the LOV sensor module in the eukaryotic kingdom. Moreover, the analysis implies that the plant circadian photoreceptors of the ZTL/ADO/FKF1-LOV family, and respectively the plant phototropins as well as the fungal WC-1 sequences, originated from two distinct endosymbiotic events that must have been separated by a few hundred million years in time.

It is interesting to note that the fungal (WC-1) and plant circadian blue-light photoreceptors (ZTL and FKF1) that both harbor a LOV domain as the light sensitive sensor switch apparently diverged from different bacterial LOV sensors and underwent convergent evolution toward the same function in plants and fungi.

Contrarily, the plant phototropins and the fungal WC-1 photoreceptors, that probably originated from the same endosymbiotic event (of an ancient α-proteobacterium)

underwent divergent evolution toward distinct functions, namely light dependent regulation of i) the circadian rhythmicity and ii) phototropic responses.
Notwithstanding, all LOV photoreceptors so far photochemically characterized (in the three kingdoms of life), apparently retained the primary photochemical events that enable photon-capture and thus (blue)-light perception.

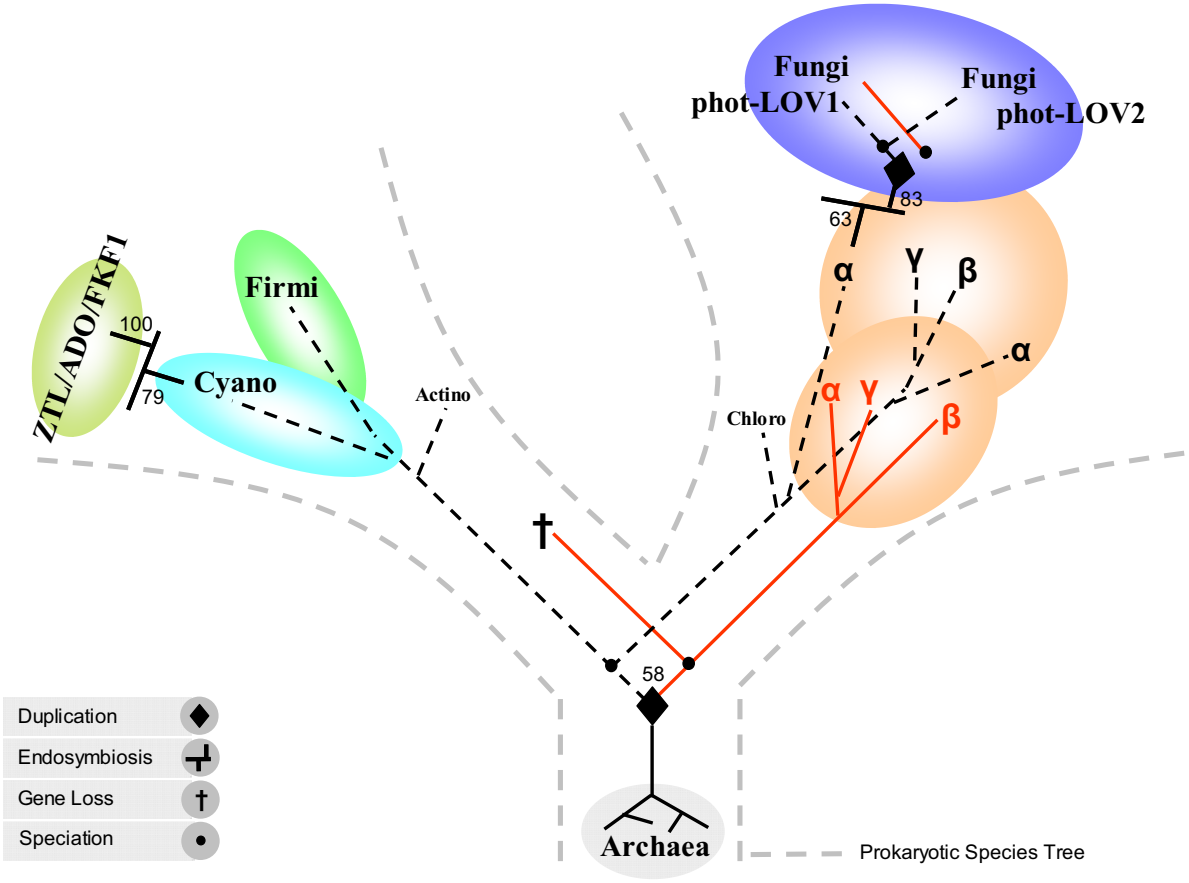


Figure 7: Evolution of LOV. The figure illustrates the evolutionary hypothesis that might have led to the wide distribution of LOV among the three kingdoms of life. The hypothesis is discussed in more detail in Chapter 2.1. In brief, an ancient duplication event after the separation of the Archaea and Bacteria may have resulted in the observed split in the phylogenetic tree, explaining the double appearance of proteobacterial LOV sequences. Subsequently a gene-loss (extinction) event has to be invoked to account for the loss of two LOV modules among the Cyanobacteria and the Firmicutes. Regarding the appearance of LOV among the eukaryotes, the affinity of eukaryotic LOV sequences to certain bacterial lineages supports the notion that all the known eukaryotic LOV photoreceptors probably originate from endosymbiosis of once free living microbes that already harbored LOV photoreceptors. Based on the observed affinities it seems likely that the appearance of phototropins and correspondingly of WC-1 in fungals coincided with the endosymbiotic uptake of an ancient proteobacterium that later became mitochondrion of the eukaryotic lineages. Respectively, the affinity of the ZTL/ADO/FKF1-LOV family to the cyanobacterial clade, suggest that the uptake of an ancient cyanobacterium that later became the eukarytic chloroplast also led to the appearance of the third photoreceptor family in plants.

Moreover, based on the presented phylogeny, LOV histidine kinase systems that are found today in a number of marine phototrophic α-proteobacteria and in a few halophilic Archaea, are suggested as among the evolutionary oldest and hence primordial LOV photoreceptors. To conclude, phototrophy (defined as a metabolic

mode in which organisms convert light energy into chemical energy for growth [101]) in ancient marine habitats might have played a primary role in the evolutionary processes that led to the appearance of the blue-light sensitive LOV-photoreceptors systems, which we find today so dispersed in the three kingdoms of life. A more detailed description of the evolutionary events that led to the widespread appearance of LOV among the pro- and eukaryotes is given in Chapter 2.1 and is moreover illustrated in Figure 7.

Functional implications derived from evolution

The phylogenetic study on pro- and eukaryotic LOV photoreceptors, which is presented in Chapter 2.1 as part of this thesis, provides insight into the evolutionary history of the LOV signaling paradigm and thus of LOV-dependent blue-light perception. But what insights can the knowledge about the (ancient) past of a given protein class provide with respect to its present function?

One example - Evolutionary constraint on the LOV dark recovery reaction?

Considering the ancestry of the bacterial LOV systems to the higher eukaryotic LOV photoreceptors that is discussed in Chapter 2.1, it seems reasonable to assume that the dark recovery in the early days of the LOV domain evolution was a solely thermally driven, and hence slow process. So the majority of the bacterial LOV photoreceptor systems should be slow (or slower compared to the plant phototropin LOV domains) in their dark recovery process. In this respect the data available for several LOV domain modules from different bacterial species seem to support this notion. Nearly all till today characterized bacterial LOV proteins are slower reverting [82, 83, 102] than their plant phot-LOV counterparts. The only exception so far is the PpSB2-LOV protein of *P. putida*, which shows a dark recovery time constant similar to the plant phot-LOV domains (see Chapters 3.1 and 5.2.1). The observation that different LOV photoreceptor classes obviously retain different dark recovery kinetics might point toward a functional importance of those different time constants for the respective LOV photoreceptor class and thus implying evolutionary constraints that appear to have governed the evolution of the dark recovery process of LOV. This hypothesis will be discussed in the next paragraph with the example of the ZTL/ADO/FKF1 LOV family.

Starting with a thermally driven dark recovery process in the early days of LOV evolution (e.g. in the bacterial world), selective (Darwinian) pressure might have caused an adaptation of the core LOV domain to assist the thermal recovery and hence accelerate the recovery process, e.g. in phot-LOV domains. Whereas, in other LOV signaling modules (e.g. the LOV domain of the circadian LOV photoreceptor family constituted by the ZTL/ADO/FKF1 proteins), apparently no such speciation toward a faster dark recovery process occurred. Hence, suggesting that the temporary fixation of the protein's light-conformation (spectrally and functionally represented by the signaling state) might have been in certain cases beneficial for the evolution of the functional response *in vivo*. Notably, recent experiments have demonstrated that for the LOV domain-containing F-box proteins ZTL and FKF1 in *Arabidopsis*, rhythmic oscillation of the protein abundance is probably regulated by direct protein-protein interaction with GIGANTEA(GI), a large plant-specific protein [103, 104]. The authors concluded that GI stabilizes ZTL and respectively FKF1 *in vivo* in a blue-light dependent manner via their LOV domains. Furthermore, mutations that result in the loss of the LOV domain's photosensitivity eliminate the blue-light-enhanced binding to GI [103, 104]. Therefore, the interaction of ZTL and correspondingly FKF1 with GI seems to occur preferentially when the LOV domain and hence the ZTL or FKF1 protein is in its light-(signaling state) conformation. Due to their slow recovery reaction, the LOV proteins (ZTL or FKF1) will be predominately in their signaling state conformation, when a (blue)-light stimulus is present (e.g. during the day). This in turn enables the interaction with GI. Whereas, in the dark a slow recovery to the LOV ground state occurs in the proteins, weakening the ZTL/FKF1-GI interaction and therefore leading to reduced photoreceptor levels. The abundance of GI is furthermore clock-controlled, so that consequently GI protein cycling confers an additional light-independent post-translational rhythm on the ZTL and FKF1 proteins, respectively. This example suggests a functional basis for an evolutionary constraint on the very slow recovery reactions that are observed for the ZTL/ADO/FKF1 LOV family and therefore exemplifies how selective (Darwinan) pressure might have controlled the evolution of the LOV domain's dark recovery process.

5.2 From signal-transduction to physiology

This chapter summarizes the results obtained in this study, regarding the signal-transduction process and the biological role of bacterial LOV domain-containing signaling systems. Additionally, unpublished data is included to strengthen the arguments presented in the published work. Due to the conserved nature of the LOV domain in bacterial and plant systems, the results obtained in this study, as well as other recent studies on bacterial LOV systems, may have implications for the understanding of the photoactivation and signal-transduction processes of LOV domain-containing signaling-systems in general. The following chapter deals with several bacterial LOV domain-containing full-length proteins or with truncated versions of the same. As a general overview, all LOV constructs relevant in this study, are depicted in Figure 8.

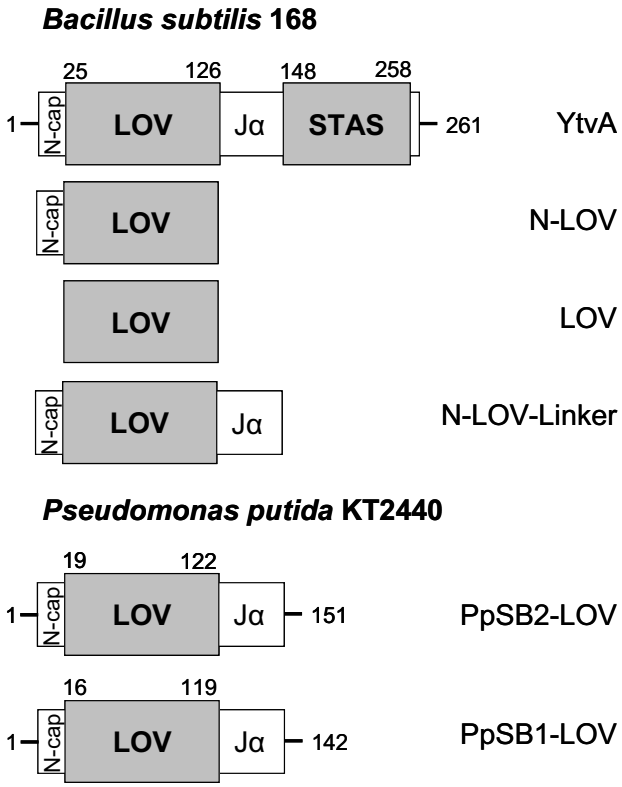


Figure 8: Relevant full-length LOV proteins and truncated LOV protein constructs investigated in this study. The *B. subtilis* YtvA protein (residues 1-261) (O34627) comprises both the N-terminal cap (N-cap), the canonical LOV domain, a probably helical linker (Jα) and the STAS domain. YtvA was expressed as N-terminal 6xHis fusion protein using pET28a in *Escherichia coli.* Truncated versions of the YtvA protein: N-LOV (1-126) containing only the N-cap and the canonical LOV domain, the isolated LOV domain (residues 25-126) and a construct comprising both N-cap, LOV domain and helical linker (N-LOV-Linker) (residues 1-147) were expressed in a similar manner to facilitate easy purification. Full length *P. putida* KT2440 LOV proteins, namely PpSB2-LOV (Q88JB0) and PpSB1-LOV (Q88E39) comprise both an N-cap, a canonical LOV domain and a probably helical C-terminal extension (Jα). The two *P. putida* LOV proteins were expressed and purified similarly to YtvA and its truncated variants.

5.2.1 The LOV photoreceptors of *Pseudomonas putida* KT2440- as a paradigm for an extended LOV construct

As discussed in Chapter 3.1, one possible mechanism that is thought to govern the signal relay from the blue-light sensitive LOV domains to the kinase moiety in the plant phot system is the unfolding of a conserved helical segment located C-terminally of the LOV2 core in plant phot [76]. Sequence analysis confirmed the presence of this helical segment (termed Jα-helix) in various plant phots, e.g. *Oryza sativa* phot1, *Zea mays* phot1, *Arabidopsis thaliana* phot1, *Pisum sativum* phot1, *Arabidopsis thaliana* phot2, *Vicia faba* phot1, *Adiantum capillus-veneris* phy3, and *Avena sativa* phot1 [77].
Between the plant phot-LOVs and the bacterial LOV domain-containing proteins this helical segment is not well conserved in sequence (Figure 9).

```
AS_PHOT1 : EHVR-------DAAERFGVMLIKKTAENIDEAAKELPDANLRPEDLWANH :  43
AC_PHOT1 : EYLEPLTKRLSQQIASFGAKIIRETAANVNEALRELPDANLKVEDLWRIH :  50
AT_PHOT1 : KHVEPVRNVIEETAVKFGEDLVKKTAVNIDEAVRELPDANMTPEDLWANH :  50
PS_PHOT1 : QHVEPLHNCIAFDTAKFGELLVKETAENVGEAVKELPDANQKPDDLWMNH :  50
VF_PHOT1 : QHVEPLHNCIAFESAKFGELLVKETAENVGEAVKELPDANQKPDDLWKNH :  50
OS_PHOT1 : EHVQ-------DDAAEFGVVLVKKTADNIDEAAKELPDANLRPEDLWANH :  43
ZM_PHOT1 : ERVR-------EAAAKDGAILVKKTADNIDEAAKELPDANLRPEDLWANH :  43
SB1      : RQVEL------ERELAFLRARPKPDERA---------------------- :  22
SB2      : AQVFA------EERVRFLEAEVAELRRQQGQAKH----------------- :  28
YtvA     : KQKE--YEKLLFDSLTFITALSTPIVPIRNGISALPLVGNLTEERFNSIV :  48
```

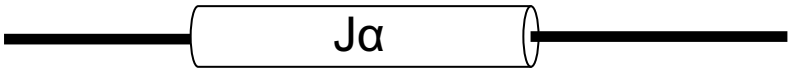


Figure 9: Multiple sequence alignment of the Jα-helix sequence regions of several phot1-LOV2 domains with the corresponding C-terminal extension of the twin PpSB1-LOV (SB1) and PpSB2-LOV (SB2) proteins. For comparison the Jα-linker of *B.subtilis* YtvA was added to the alignment. Sequence abbreviations of the eukaryotic taxa from which the phot1-LOV2-Jα sequence was used in the alignment are as follows: AS: *Avena sativa*, AC: *Adiatum capillus-veneris*, AT: *Arabidopsis thaliana*, PS: *Pisum sativum*, VF: *Vicia faba*, ZM: *Zea mays*).

Nevertheless, the presence of a helical linker in *B. subtilis* YtvA (for details see Chapter 3.2) located C-terminally to the LOV core, thus connecting the LOV and STAS domains has been proposed [105] based on UV-circular dichroism (CD) - spectroscopy and secondary structure predictions (see Chapter 3.2). Notably, its presence could very recently been proven by the determination of the crystal structure of an extended YtvA-LOV construct, which included the prospective linker polypeptide (residues 127-147) [106].

In Chapters 3.1 and 3.3 the cloning and characterization of two novel LOV domain-containing proteins from the rhizosphere colonizing organism *P. putida* KT2440 is described.

The proteins PpSB1-LOV (Q88E39) and PpSB2-LOV (Q88JB0) contain a canonical LOV core as well as short C- and N- terminal extensions (Figure 10). The two probably paralogous proteins are highly similar in sequence as they share about 66% identical positions. Interestingly, both lack a fused output domain, but bioinformatic analysis of inter-organismal operon organizations demonstrate a clustering of a helix-turn-helix transcriptional regulators as the putative LOV co-localized (in close proximity, but not fused) effector domains.

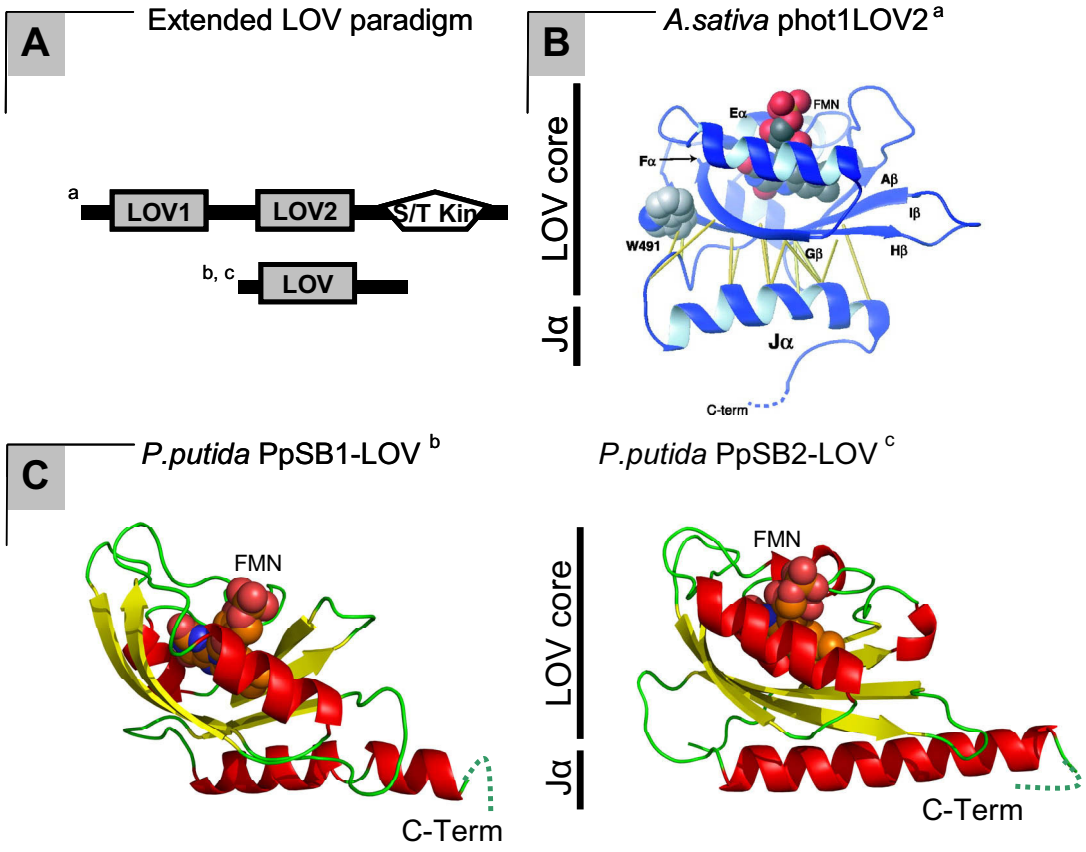


Figure 10: Illustration of the extended LOV paradigm. A) Full-length domain architectures of plant phototropins (a) and of the twin LOV proteins of *P. putida* KT2440 (b,c). Whereas, plant phots contain two LOV domains (LOV1 and LOV) fused to a Serine/Threonine Kinase (S/T Kin), the two *P. putida* proteins consist only of a core LOV domain with short N- and C-terminal helical extensions. B) Structural model of the extended LOV-construct of Asphot1 LOV2 as derived from NMR spectroscopy. C) Homology models of the *P. putida* LOV proteins (PpSB1-LOV and PpSB2-LOV) with the C-terminal helical extension (Jα-helix) computationally docked to the core LOV domain (see Chapter 3.1 for details). In both Asphot1 LOV2(B) and in the twin LOV proteins of *P.putida* (C), the Jα-helix pachs/docks well against the central β-sheet surface of the core LOV domain. Figure 10B was adapted from [76] and is reprinted with permission of the American Association for the Advancement of Science (AAAS).

Secondary structure predictions as well as UV-CD experiments indicate that in both proteins the LOV core extensions (N- and C-terminal) might be helical (for details see later this chapter). Taking this information into account it was possible to generate a structural model for PpSB2-LOV and respectively for PpSB1-LOV (data not published) that comprises the C-terminal helical extension to the LOV core. The C-terminal helical segment could be computationally docked well underneath the central β-sheet surface of the PpSB2-and PpSB1 LOV domain (Figure 10).
This renders the two proteins of *P. putida* PpSB1-LOV and PpSB2-LOV naturally occurring paradigms for an extended LOV construct, very similar in structure to the recombinant construct of Asphot1-LOV2, studied by Harper and colleagues by NMR spectroscopy [76] (see Figure 10).

PpSB1-LOV and PpSB2-LOV exhibit a conserved phot-like photochemistry

The two duplicated proteins of *P. putida* exhibit a phot-like photochemistry as demonstrated in Chapter 3.3 for PpSB1-LOV and in Chapter 3.1 for PpSB2-LOV, respectively. Interestingly, the highly conserved canonical sequence motif of LOV domains, GXNCRFLQG, is found to be changed to YQDCRFLQG in both of the *P. putida* proteins and is moreover conserved among all saprotrophic Pseudomonads that contain similar proteins. Notably, this exchange does not inhibit or alter the FMN binding affinity as well as the phot-photochemistry in the case of the two proteins studied as part of this thesis. In contrast to the other bacterial LOV domain-containing proteins characterized so far, the dark state recovery of PpSB2-LOV was comparably fast (τ_{REC}= 114 s at 20°C) and hence more similar to plant phot LOV domains (e.g. *C. reinhardtii* LOV1, τ_{REC}=334 s at 20°C). On the other hand its paralogous twin protein PpSB1-LOV shows the slowest recovery rate ever observed among bacterial LOV domains (τ_{REC}= 29.5 h at 20°C). The recovery kinetics for PpSB1-LOV therefore more closely resembles values observed for the plant FKF1-LOV domain [67].
Interestingly this equips *P. putida* and some other Pseudomonads with one LOV system (PpSB1-LOV) that will in the light be predominantly “switched on” (due to the accumulation of the signaling state because of its very slow recovery). Additionally the organism possesses a LOV photosensor system (PpSB2-LOV) whose recovery is so fast that it will return nearly instantaneously to the ground state and hence will more easily be “switched off”. The functional relevance of this remarkable dichotomy, if not an evolutionary artifact, will be a very interesting point of future studies. Solely for comparison *B. subtilis* YtvA (τ_{REC}= 2600 s at 20°C) and its isolated

LOV domain YtvA-LOV (τ_{REC}=3900s at 20°C) [102] show intermediate values in between those of PpSB1-LOV and PpSB2-LOV. The structural basis of the observed variability of those dark recovery time constants, especially among the bacterial LOV domain-containing proteins is not yet well understood.

However, as discussed in Chapter 3.3, several mechanisms were proposed that may influence the dark recovery rate in LOV domains. One study pointed towards a general base catalyzed recovery reaction, probably driven from two surface exposed histidine residues in the *Avena sativa* LOV2 domain [107], thus confirming an earlier hypothesis of Swartz *et al.* (2001) [64] that suggested a base catalysis controlling the dark recovery in LOV domains. It is thus concluded that base catalysis in at least Asphot1-LOV2 can take place from those two histidine residues although they are located in about 12 Å distance from the FMN isoalloxazine ring, via a hydrogen bonding network of the base, chromophore, and intraprotein water molecules [64]. In case of the slow reverting PpSB1-LOV and correspondingly the fast reverting PpSB2-LOV, the presence of four histidines in PpSB2-LOV (with one of those histidines, His64, probably in close proximity to the flavin isoalloxazine ring) and correspondingly the complete absence of histidine residues in PpSB1-LOV might indeed suggest the involvement of similar base catalyzed mechanisms as the one postulated for Asphot1-LOV2.
However, as other slow reverting bacterial LOV domains and LOV domain-containing proteins (e.g. *B. subtilis* YtvA) as well as the very slow reverting plant LOV domains of the ZTL/ADO/FKF1 family, possess histidine residues throughout the LOV core it seems rather unlikely that the recovery process is solely governed by a histidine-driven base catalyzed mechanism. It should be noted that the light-sensitive LOV domains of the ZTL/ADO/FKF1 photoreceptor family, possess an 11 aa insertion between the αE and αF-helices which is absent in other LOV domains (comparable to VIVID-LOV in the alignment in Figure 4, Chapter 1.3 of the introduction) [17]. The additional segment, probably adopting a random coiled conformation, might contribute to the slower recovery kinetics that are observed in this LOV-photoreceptor family. The additional segment might thus provide more conformational flexibility in the LOV domains signaling state and, therefore, alleviate strain in the protein associated with the FMN-protein linkage. This in turn might slow down the recovery reaction.

A second proposal recently suggested an additional mechanism based on the following: in a random mutagenesis study conducted by Christie *et al.* (2007) on the LOV2 domain of Asphot1, an isoleucine residue located within van der Waals distance from the highly conserved photoactive cysteine was identified to be involved in the dark recovery reaction. It was suggest by the authors that this isoleucine probably makes contact with its methyl-group to the sulfhydryl group of the cysteine [108] thus creating strain in the protein to stabilize the signaling state intermediate. Mutation of this isoleucine to valine or leucine resulted in a dark recovery that was faster by one order of magnitude, or about three fold (for the isoleucine to leucine mutation), compared to the wildtype protein.

In the case of PpSB1-LOV and PpSB2-LOV, the position of the isoleucine residue identified in Asphot1-LOV2 as hot-spot that strongly influences the dark recovery reaction, is occupied by a leucine, which might account for the relatively fast recovery reaction observed for PpSB2-LOV. Contradictory to this reasoning, in both *P. putida* proteins that behave so dramatically different with respect to their recovery kinetics, the identified hot-spot protein region is highly conserved in sequence, thus *a priori* suggesting that the different recovery kinetics can not depend largely on this position in the protein.

In conclusion, both mechanisms elaborated in the above paragraph might contribute to some extend to the dark recovery reactions also in bacterial LOV proteins. However, as both hypothesis fail to unequivocally explain the different recovery kinetics of PpSB1-LOV and PpSB2-LOV, other additional mechanisms have to be considered that influence the recovery kinetics in the bacterial LOV domain family.

The dramatic differences in the dark recovery kinetics observed for PpSB1-LOV and PpSB2-LOV, together with the high degree of sequence conservation should render this system an ideal starting point to investigate the mechanisms underlying the currently not well understood dark recovery reaction of LOV domains by means of directed and/or random mutagenesis.

Tryptophan fluorescence suggest conformational changes in the C-terminal region of PpSB2-LOV

Tryptophan (Trp) fluorescence spectroscopy on W97, the sole Trp residue in PpSB2-LOV, revealed a transient increase of the Trp-fluorescence in the protein's signaling state that was, although low in magnitude, completely reversible in the dark with a

time constant (τ_{RECW}= 102 s) quite in accordance to the observed dark recovery time of the protein. This suggests conformational changes in the microenvironment of the this Trp residue, larger than in the *B. subtilis* YtvA protein where changes in Trp fluorescence associated with the signaling state formation are minor. [109].

The conserved Trp of PpSB2-LOV (W97) is probably located in a similar environment as in the NMR derived structure of the extended LOV2 construct of Asphot1, namely in the vicinity of the loop that connects the LOV core to the helical extension, hence the observed changes in the Trp fluorescence upon signaling state formation as well as the presence of a similar C-terminal helical segment in both PpSB1-LOV and PpSB2-LOV might point towards a common signal-transduction mechanism between the plant phot system and the two twin proteins of *P. putida*.

UV-CD spectroscopy to study the solution structure of PpSB2-LOV (unpublished)

In order to test this hypothesis and evaluate the low-resolution 3-dimensional model for the two proteins, UV-CD spectroscopy was employed. In principal this method should allow to follow light driven conformational changes in PpSB1-LOV and PpSB2-LOV that might accompany the signaling state formation, e.g. an unfolding of the C-terminal extension. However, one has to keep in mind that transient changes that occur below a timescale of seconds as well as 3-dimensional rearrangements of protein secondary structure elements (e.g. movement of helices without unfolding) that are not accompanied by a global change in the secondary structure content are not resolvable with this approach. Figure 11 shows the recorded UV-CD spectra of PpSB1-LOV and PpSB2-LOV, respectively.

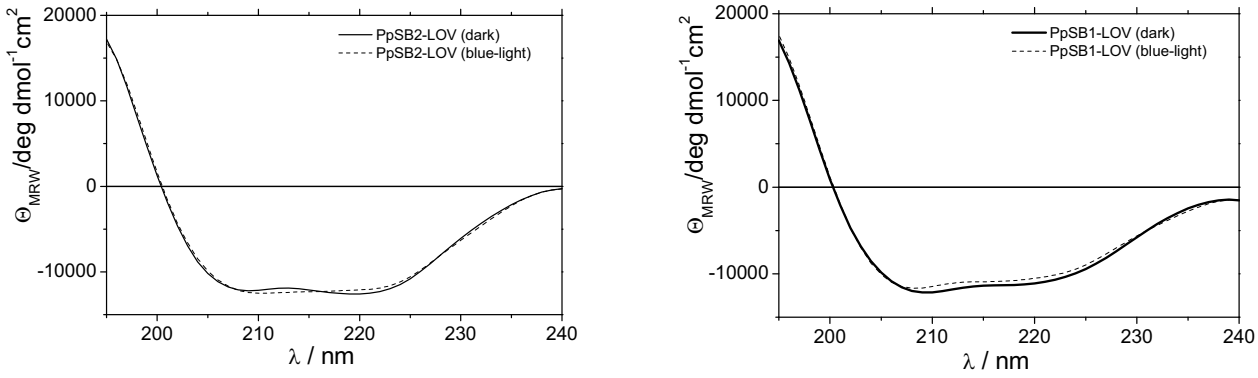


Figure 11: CD spectra in the UV region recorded for the dark (dashed) and light (signaling) state (dots) of PpSB1-LOV (left) and PpSB2-LOV (right). Experiments were performed in co-operation with Aba Losi and Valentina Buttani at the University of Parma, Italy. Experimental conditions were the same as described in Chapter 3.2.

Spectra were recorded before (dark state) and after illumination with blue-light (signaling state). Table 1 summarizes the data derived from a CCA-analysis of the PpSB1-LOV and PpSB2-LOV CD-spectra.

TABLE 1: Results of the CCA analysis on CD spectra

Sec. structure	% SB2-LOV #(168 aa)		% SB1-LOV §(162 aa)	
	dark	light	dark	Light
I. α-helix	33.4 $\pm$ 4.9 (56 $\pm$ 8)	33.4 $\pm$ 2.5 (56 $\pm$ 4)	36.8 (60)	40.5 (66)
II. RC	19.1 $\pm$ 4.7 (32 $\pm$ 8)	18.5 $\pm$ 5.5 (31 $\pm$ 9)	16.7 (27)	20.0 (32)
III. β-turns/others	20.5 $\pm$ 4.0 (34 $\pm$ 7)	20.8 $\pm$ 4.7 (35 $\pm$ 8)	23.4 (38)	25.0 (41)
IV. $\beta_{twisted}/\beta_{Par.}$	9.8 $\pm$ 5 (16 $\pm$ 8)	9.7 $\pm$ 4.3 (16 $\pm$ 7)	10 (16)	10.8 (17)
V. β Antipar.	17.2 $\pm$ 1 (28 $\pm$ 2)	17.6 $\pm$ 1 (29 $\pm$ 2)	10.1 (16)	3.7 (6)
$^{\S}\sum_{i=1}^{n}[y_i - f(\lambda)]^2$	*26 $\pm$ 12*	*15 $\pm$ 8*	*8.5*	*7.3*

†: The statistical error is the standard deviation and derives from two sets of measurements on two different preparations for PpSB2-LOV and N-LOV-Linker.
§: Average squared error, where y_i = experimental curve, $f(\lambda)$ = fitting curve;
‡: The preparations of PpSB2-LOV and PpSB1-LOV contained considerable amounts of apo-protein. The CCA-algorithm and its application to analyze the CD-data in Figure 10 is described in [105].

In accordance with the CD data presented for the isolated LOV domain of YtvA (see [105], Chapter 3.2 and the discussion in Chapter 5.2.2) and furthermore under the assumption of a canonical LOV fold present in PpSB1-LOV and PpSB2-LOV, the two additional elements, located N- and C-terminally to the core LOV domain, should both be helical in conformation.

UV-CD spectra indicate no global conformational changes in the two proteins upon illumination, hence a rearrangement in PpSB2-LOV as implied by changes in the Trp-fluorescence is apparently not caused by an unfolding or dissociation of the C-terminal extension as it was suggested for Asphot1-LOV2 [76]. It should be noted

that the CD-data presented here are still associated with a relatively large experimental error (about 15% for PpSB2-LOV), probably attributed to the presence of apo-protein in the preparations of PpSB1-LOV and PpSB2-LOV. Therefore, an optimization of the PpSB1- and PpSB2-LOV expression that increases the chromophore content in the LOV protein should reduce the experimental error and hence provide more reliable data. Nevertheless, UV-CD-spectroscopy implies that the N- and C-terminal extensions in the two *P. putida* proteins are helical in structure, thus at least suggesting similar structural elements located outside the canonical core LOV domain in both the bacterial and eukaryotic phot-LOV systems.

In conclusion, although PpSB1-LOV and PpSB2-LOV probably resemble structurally the extended construct of Asphot1-LOV2 that was studied by Harper and co-workers using NMR spectroscopy [76], similar signal-transduction processes for the bacterial and eukaryotic LOV systems, namely an unfolding of the C-terminal extension (Jα-helix) outside the core LOV domain could not be detected with the presented approach. Therefore, currently all available data (see also the following chapter for a discussion) points toward different light activated signal-transduction systems in the plant and bacterial LOV photoreceptors, although structural elements (like N-cap and Jα-helix) seem to be conserved between the two kingdoms of life.

5.2.2 LOV – LOV dimerization and interdomain interactions in the *Bacillus subtilis* YtvA protein

Proteins of the structurally conserved PAS fold, to which the LOV domain family of signaling modules belongs, are often involved in general sensory functions, transcriptional activation and regulation, both mediated by ligand binding, direct protein-protein interactions and homo/hetero-dimerization [88]. Currently the information regarding the oligomerization state of LOV domains and / or of LOV domain-containing proteins is confusing, especially with respect to the physiological significance of those processes for the signal-relay to downstream partners. In Chapter 3.2 a possible LOV-LOV dimer complex formed by the isolated LOV domain of the *B. subtilis* YtvA protein is described. Based on size-exclusion chromatography studies, CD spectroscopy and molecular docking simulations a common surface for LOV-LOV and intra-protein interaction was proposed for full-length YtvA.
As outlined in Chapter 3.2, the oligomerization state of different LOV domains is still controversially under debate. Furthermore, it is not unequivocally resolved for the LOV constructs that clearly dimerize in solution such as Asphot1 LOV1, the LOV1 domain of Atphot1 and Atphot2 as well as the LOV domain of FKF1 of *Arabidopsis*, how the corresponding LOV monomers orient towards each other to form the dimer complex. This raised the question, whether there is a general dimerization / interaction surface in some LOV domains that is absent in others e.g. in Asphot1 LOV2, that appeared to be monomeric.

A common surface in the LOV domain might govern the LOV-LOV and LOV-STAS interactions in YtvA

In Chapter 3.2 it was shown that the isolated LOV domain (LOV, residues 25-126) of YtvA forms a stable dimer in solution, whereas full-length YtvA is largely monomeric. This strongly suggests a common interface for LOV-LOV dimerization and for the LOV-STAS interaction, respectively. Furthermore, the dimerization of the LOV domain is not prevented by the presence of the N-terminal cap (N-LOV, residues 1-126) and the Jα-linker (Losi, personal communication).

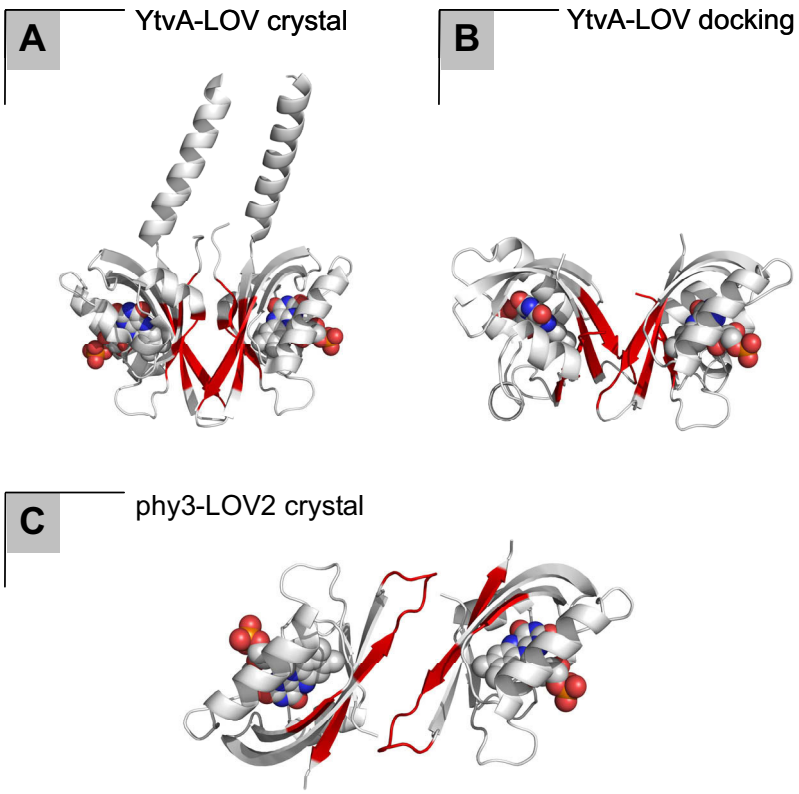


Figure 12: Illustration of the possible dimer complexes observed for the LOV domain of different LOV-photoreceptor proteins. A) Head to head dimer complex observed in the crystal structure of YtvA-LOV [106] (pdb-entry: 2PR5). The N-terminal Jα-helix protrudes from the core LOV domain and does not pack against its β-sheet surface. B) In comparison the ZDOCK derived computationally docked complex of YtvA-LOV (lacking the Jα-helix), as described in Chapter 3.2 [105], is depicted. C) Dimer complex observed in the crystal-structure of phy3-LOV2 showing a different orientation of the LOV core monomers as in the YtvA-LOV dimer (A and B). Interaction hot-spots predicted using the computational alanine scanning tool (Robetta software suite [110]) are highlighted in red in the corresponding structures and models. The figure nicely illustrates that the same interaction surface, namely the central β-sheet surface, is probably involved in the dimerization process.

Applying UV-CD spectroscopy to study the secondary structure composition of the three constructs (LOV, N-LOV and full-length YtvA), it could be demonstrated that both the N-terminal cap as well as probably the linker region (aa 127-147) are mostly helical. Furthermore, it was possible to computationally generate YtvA-LOV dimer complexes by employing a molecular docking approach. In the most favored docking complex the LOV monomers form a head-to-head dimer facing each other with the central β-sheet surface (Figure 12). The proposed complex might be stabilized by hydrophobic and electrostatic interactions, whereas a cluster of hydrophobic amino acids at the LOV-LOV interface would nicely account for the stability of the dimer in solution. It should be mentioned here, solely for comparison, that the twin LOV proteins of *P. putida* both form stable dimers in solution (see Chapter 3.1 for PpSB2-

LOV gel-filtration experiments; corresponding experiments were also conducted for PpSB1-LOV, revealing a dimeric organization of the protein *in vitro* (unpublished observation)).

Notably, also in the only recently solved crystal structure of YtvA-LOV, that comprised the canonical LOV domain as well as the proposed helical linker (residues 20-147), two LOV monomers form a head-to-head dimer complex [106] similar to the one observed in the docking approach described in Chapter 3.2. In the crystal, the two LOV monomers face each other with their central β-sheet surfaces, whereas the clearly alpha-helical linker (now termed Jα-helix), protrudes from the LOV core making intermolecular contacts with the corresponding alpha helices of symmetry related molecules (Figure 12). However, in solution, this might not be the case, unpublished CD data by Losi and co-workers indicates that in a construct that comprises additionally to the core LOV domain of YtvA also the short N-terminal cap as well as the Jα-helix (N-LOV-linker, see Chapter 5.2), the Jα-helix is disordered, whereas the proteins retains its dimeric organization (Losi, personal communication).

Both observations regarding the Jα-helix in YtvA, whether the protruding Jα in the crystal [106] or the disordered Jα in solution (Losi, personal communication), seem contrary to the situation in the extended LOV2 construct of Asphot1 that was previously studied by solution NMR-spectroscopy. In this construct, the Jα-helix neatly packed against the core LOV domain in the dark [76].

Sophisticated methods such as time resolved thermal grating (TG) and transient thermal lensing (TrL) techniques have recently been applied to determine the change in the molecular diffusion coefficient (D) and hence assess the oligomerization state as well as conformational changes upon photoexcitation of several LOV domain constructs in a time resolved manner [111, 112]. Such experiments with the isolated Atphot1 LOV2 domain (lacking the Jα-helix) indicated that the protein is monomeric in the dark for concentrations < 100 µM. However, a transient decrease of D can be monitored following the signaling state formation, which was interpreted by the authors in terms of dimerization of Atphot1-LOV2 [112]. On the other hand, when protein concentrations were kept above 100 µM, the protein was dimeric in the dark state and underwent a light-induced dissociation with a time constant of 300 µs.

Similar experiments using an extended Atphot1-LOV2-linker construct that included the C-terminal Jα-helix, suggested that the very same LOV2 domain, in the presence of the linker (Jα-helix) is monomeric in the dark state. Furthermore, the authors

demonstrated a light induced decrease in D for the LOV2-linker construct with a time constant of 1 ms, being interpreted as unfolding of the Jα-helix. Transient thermal lensing (TrL) experiments furthermore suggested that this unfolding is preceded by a dissociation of the linker from the LOV core with a time constant of 300µs [113].
This implies that the presence of the linker and probably its hydrophobic interactions with the core LOV2 domain inhibit the previously observed LOV-LOV dimerization that was probably mediated by the same hydrophobic surface on the LOV domain side. Similar time constants for the dissociation of the Atphot1 LOV2-dimer (~300µs) and the linker dissociation from the Atphot1 LOV2-linker construct (~300µs) moreover suggested that the conformational changes upon signaling state formation cause a change in the hydrophobic interactions between either two LOV2 monomers, or between the LOV2-core and the Jα-helix, respectively.

The available experimental data regarding the oligomerization of LOV domains (both from bacterial and eukaryotic LOVs) seems confusing and even in some cases contradictory at the first glance. Nevertheless, the results described in Chapter 3.2 regarding the oligomerization state of full-length YtvA and its truncated LOV domain constructs, the observed dimer complex in the recently solved crystal structure of YtvA-LOV as well as the latter described TG and TrL experiments on Atphot1 LOV2 and its extended construct, strongly highlight the importance of LOV oligomerization and moreover of the C-terminal linker region (Jα-helix) in the LOV signal-transduction process *in vitro*. However, the biological significance of the transient conformational changes in phot-LOV domains has yet to be demonstrated.

Implications for the signal-transduction mechanisms of YtvA and the LOV domain signal-relay

In conclusion, these very different experiments, although they still lack an *in vivo* link to physiological responses, emphasize all one important aspect: the central β-scaffold might be the common interaction surface involved in the LOV domain mediated signal-transduction process. Whereas the blue-light signal that is initially received by the FMN-chromophore in the LOV domain, is subsequently relayed via global conformational changes probably involving the central β-sheet surface of the LOV core, to the coupled effector domain(s). In addition, the C-terminal linker (or Jα-helix) probably strongly influences this signal-relay as it might act as a competitor for

LOV-LOV homo/hetero-dimerization in full length LOV proteins (e.g. plant phots or YtvA) or might act as a competitor for LOV-effector (e.g. LOV-kinase or LOV-STAS) interactions by blocking the common interaction surface. Under these assumptions the dissociation / unfolding of the Jα-helix and hence the loss of its amphipathic character might expose the central β-sheet, rendering it accessible for an interaction with e.g the kinase moiety in full-length phot or the STAS domain in full-length YtvA. However, one has to keep in mind that in the crystal structure of YtvA the Jα-helix does not pack against the central β-sheet. Furthermore, no unequivocal unfolding or dissociation of the Jα-helix in YtvA could so far be observed (see Chapter 3.2) and hence the model described her for Asphot1 LOV2 might not be altogether relevant for YtvA.

This phot-LOV2 activation model, described in the previous paragraph, was as well challenged by the recent observation that the Atphot2 LOV2 Jα-helix is not needed for the interaction between LOV2 and the kinase domain and furthermore is dispensable for the light activated phosphorylation reaction of a heterologous substrate [114]. More importantly, it is apparently not needed for the “repression” of the reaction in the dark. Under the aforementioned assumptions one would have to expect that the absence of the linker, resulting in the constitutive exposure of the β-sheet surface, should trigger constitutive kinase activation even in the absence of light. However, as this is apparently not the case (at least in the *in vitro* assay system used by Matsuoka and co-workers [114]), two possible interpretations remain. Firstly, light-dependent autophosphorylation of phototropin might proceed via a different molecular route than the substrate phosphorylation by phototropin that was used to assay kinase activity in their study [114]. Secondly, light-driven conformational changes on the LOV2 core domain itself, additionally to the unfolding and decoupling of the Jα-helix from LOV2, might be necessary to mediate autophosphorylation of the kinase and/or substrate phosphorylation by phototropin, respectively.

However, as already stressed, concerning YtvA one has to keep in mind that no unfolding of the Jα-helix could by observed for full-length YtvA or its truncated constructs by using UV-CD spectroscopy as described in Chapter 3.2. In contrast and contradictory to the results described in Chapter 3.2, Möglich *et al.* (2007) [106] observed an about 10% decrease in helical content in YtvA-LOV upon illumination, when the protein was studied by UV-CD spectroscopy in solution. However, as a

similar effect was not observed in the light-state crystal-structure of YtvA-LOV, these results have to be interpreted with great care.

In conclusion, based on the currently available data, different photoactivated signal-transduction mechanisms have to be considered for the plant phot system and the *B. subtilis* YtvA protein. Nevertheless, the signal-transduction process might involve a similar interaction surface on the LOV domain, namely the central β-scaffold as well as a helical linker (Jα-helix) located outside of the canonical LOV core in both the bacterial and the eukaryotic phot systems. Moreover, contrary to the phot-LOV system, all experimental evidence currently available for *B. subtilis* YtvA, points toward a direct interaction between the LOV and STAS domain of the full-length protein (Losi, personal communication). Nevertheless, the Jα-linker may still be necessary for the signal-transduction process in YtvA, but might play altogether a different role as in phot-LOV2.

With the advent of functional assays to probe light driven physiological responses, for example in *B. subtilis* and involving its LOV blue-light receptor YtvA, mutational studies in the future should provide a promising experimental approach to probe interaction surfaces and thus elucidate the signal-transduction mechanism of the YtvA protein of *B. subtilis*.

5.2.3 Biological role of bacterial LOV domain-containing proteins

In Chapter 3.3, two LOV signaling modules, found in the saprotrophic γ-proteobacterium *P. putida* KT2440, are described. The paralogous twin proteins lack both a fused effector domain but possess short N-terminal and C-terminal extensions that are probably helical as already described in Chapter 5.2.1. Based on bioinformatic analyses it was suggested that helix-turn-helix (HTH) transcriptional regulators might be the possible output domains of the LOV initiated signaling cascade. This hypothesis is based on the conservation of operon-like structures in the genomes of the analyzed Pseudomonds. In those gene clusters the corresponding LOV protein is always co-localized with two HTH-transcriptional regulators. Furthermore, the occurrence of LOV-HTH fusion proteins among several bacterial species suggests at least the possibility of a signal-relay between the two partners in a two component-signal-transduction system.

More extensive analysis of the genomic context (neighboring genes) of the two LOV proteins in *P. putida* KT2440 showed a clustering of genes annotated to be involved in the iron-limitation / biofilm development and/or stationary phase cell lysis response in the organism. For example several ferric-iron and iron-siderophore receptors cluster together with PpSB1-LOV, implying a regulatory role in the response towards severe iron-limitation. Moreover, a dicistronic operon consisting of two genes LrgA/B is highly conserved throughout all *Pseudomonas* specific LOV gene regions. In *Staphylococcus aureus* the LrgA/B system, together with its homologs cidA/B, is involved in the regulation of mediated cell death and lysis [115]. Moreover, recent studies suggested their involvement in the development and adherence of biofilms of the organism [116]. This latter process is intimately linked In Pseudomonads with iron-acquisition and consequently with the production of siderophores (e.g. pyoverdine) [117]. Pyoverdines (PVDs) are a structurally conserved group of siderophores (iron-scavengers) that are produced by various fluorescent *Pseudomonas* species, such as the pathogen *P. aeruginosa* [118], the saprotrophic (root colonizing) *P. putida* [119], and the plant pathogen *P. syringea pv. syringae* [120]. In the majority of fluorescent *Pseudomonas* strains, PVDs represent the primary iron-acquisition system, although some species can produce additional siderophores such as pyochelin and quinolobactin [121]. The structure of pyoverdines can vary extremely among species and even between strains of the

same species [117]. As one example, the chemical structure of the pyoverdine from *P. aeruginosa* ATCC 27853[122] is depicted in Appendix B. The interest in PVDs has considerably increased in recent years, as it became apparent that these compounds play several important roles in *Pseudomonas* biology [117]. PVD mediated effects range from control of infection and virulence to cell-to-cell signaling in *P. aeruginosa* [123], plant-growth promotion and phytopathogen biocontrol by PVD producing *P. putida* and *P. fluorescens* species [124, 125], to a possible role in mediating phytopathogenesis of plant pathogens such as *P. syringae* [126].

Taken together, both inter-organismal gene-cluster comparison as well as genome proximity analyses imply a regulatory role in the iron-starvation response for the two putative LOV photosensors, therefore, suggesting more generally the importance of blue-light as a stimulus under severe iron-limitation in *P. putida* KT2440.
In order to test this hypothesis, several experiments were performed to study the influence of light in iron-limitation related processes. Most notably, the iron-siderophore pyoverdine was secreted to the culture supernatant in about 2-fold higher levels when *P. putida* was grown under continuous blue-light illumination compared to values determined for the same strain grown in the dark. Such an increase in the pyoverdine secretion was not observed in supernatants of red- and green-light grown cultures. This strongly suggests the involvement of a blue-light photoreceptor in this process. An initial characterization, using the crystal-violet staining assay described by O'Toole and Kolter [127], to elucidate the influence of blue-light on the biofilm formation, indicated a slight increase in the biofilm biomass when the strain was grown under blue-light compared to darkness grown cells. However, it should be noted that the experimental error associated with this assay system is quite large and more sophisticated methods, such as confocal scanning microscopy of statically or flow-chamber grown biofilms should be used in the future to evaluate the influence of blue-light on the formation and development of biofilm in *P. putida* in greater detail.

The central question arising from the sheer abundance of blue-light photoreceptor modules among non-photosynthetic prokaryotes is the following: Why should blue-light be an important environmental stimulus for non-photosynthetic bacteria (no matter if beneficial or harmful)? Based on the observations discussed in Chapter 3.3

a hypothesis can be brought forward that might explain the biological necessity for *P. putida* and probably other Pseudomonads, to sense blue-light under severe iron-limitation. In general, this scenario would probably be applicable for most prokaryotes as will be elaborated in more detail in the following paragraph.

A possible scenario for the blue-light dependency of the iron starvation response in P. putida and other prokaryotes

Heme (an iron containing porphyrin-derivative, Figure 13) is the prosthetic group in many proteins that are involved in a myriad of cellular processes such as respiration, O_2 transport, electron-transport, oxidative stress responses, O_2-sensing, oxygenation reactions and detoxification reactions (see [128] and references therein). Furthermore, heme is ubiquitous in probably all prokaryotes [128]. Its biosynthesis pathway is concluded with the incorporation of ferrous-iron into the porphyrin-ring system of the heme precursor protophorphyrin IX (PPIX). This step is catalyzed by an enzyme named ferrochelatase (hemH, protoheme ferrolyase, E.C. 4.99.1.1) (Figure 13), which is ubiquitous from bacteria to man [128].

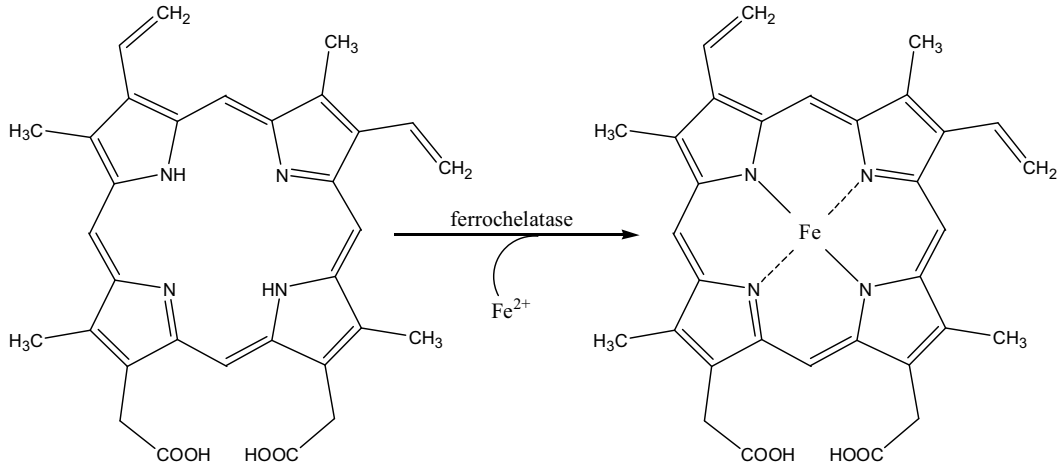


Figure 13: Final catalytic step of the heme-biosynthesis pathway. Heme is produced by the action of a ferrochelatase, which incorporates ferrous-iron (Fe^{2+}) into the porphyrin ring system of protoporphyrin IX (PPIX).

This catalytic step might thus represent a bottleneck in the whole biosynthesis pathway, as the lack of iron should ultimately result in the accumulation of PPIX in the cell [129]. It was shown for *E. coli* [130] that inactivation of *hemH* results in a rapid killing of the mutant strain under visible light, probably due to the cellular accumulation of PPIX. In analogy, a similar mechanism was recently proposed for *P.*

fluorescens by Baysse *et al.* [131]. Consequently, concluding that PPIX must be the causative agent of the observed phototoxicity effect. This effect is probably related to the light dependent formation of reactive oxygen species (ROS) from the porphyrin ring system of PPIX. Porphyrins in general can be excited by blue-light with a high yield to the corresponding triplet state, which subsequently reacts forming free radicals. Eventually, those radical species give rise to ROS, in a process involving energy transfer with molecular oxygen. ROS are known to cause oxidative damage to many cellular tissues and organs [42].

A similar principle is utilized in medicine for photodynamic therapy (PDT), which is applied in the clinical treatment of skin cancers and various skin diseases [132]. In this treatment, usually 5-aminolevulinic acid (ALA), the product of the rate-limiting step in the porphyrin biosynthesis, is administered. This results in an increase of cellular concentrations of PPIX, which is more pronounced in cancerous cells than in healthy tissue [133]. PPIX in turn acts as photosensitizer inducing the formation of singlet oxygen (ROS) and thus consequently leading to cell death [42] under illumination.

In summary, in our bacterial setting, lack of iron and thus accumulation of PPIX with a simultaneous exposure to blue-light should cause at a minimum an oxidative stress to the bacterial cells and might in the extreme result in conditional cell death.

Therefore, it seems feasible for both pro- and eukaryotic species to develop systems to cope with the harmful effects of (blue)-light induced oxidative damage. A summary of the possible bacterial response mechanisms is depicted in Figure 14.

The easiest answer would of course be to escape from those harmful light conditions, which is performed in nature via microbial phototaxis [134]. However, light avoidance responses might not always be applicable as in the case of non-motile microorganisms or might not be possible, e.g. in certain environments. Therefore other mechanisms had obviously to be developed in response. For example, some non-photosynthetic bacteria such as *Myxococcus xanthus*, *Pantoea agglomerans* (formerly *Erwinia herbicola*) and *Streptomyces coelicolor,* as well as many photosynthetic prokaryotes, such as *Rhodobacter capsulatus*, *Synechococcus* and *Anabaena* strains produce carotenoids (see [135] and references therein) as protective pigment against light induced oxidative-damage [135]. Similarly, some photosynthetic species produce carotenoids as light-harvesting and triple-quenching, protective pigments to prevent chlorophylls and bacteriochlorophylls to be exposed to

harmful photodestructive reactions in the presence of oxygen [136]. In those prokaryotic organisms as well as in certain eukaryotes (e.g. in the fungus *Neurospora crassa* [94]), the carotenoid synthesis is directly regulated by blue-light. Besides carotenoids, other pigments like for example melanins [137] might be utilized in this light protective response.

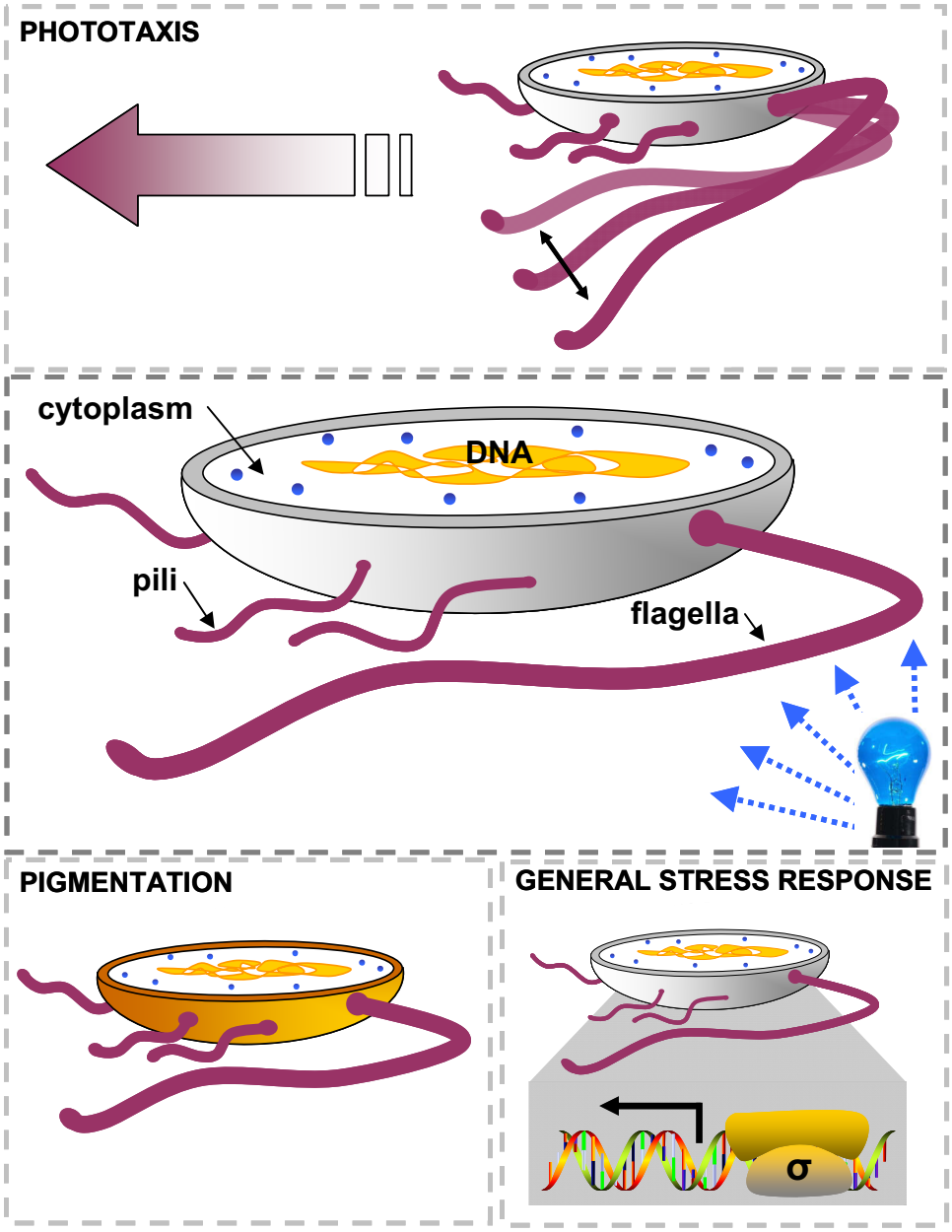


Figure 14: Illustration of possible bacterial blue-light response mechanisms. Blue-light might induce the formation of reactive oxygen species (ROS) in bacterial cells, that contain endogenous photosensitizers such as porphyrins, flavins etc. Therefore, a bacterial cell, when exposed to blue-light irradiation, might response with i) phototactic avoidance (by pili or flagellar motility), ii) production of protecting pigments, such as carotenoids and melanins or iii) with the induction of the general/oxidative stress response regulon (by an alternative sigma factor), to cope with the harmful effects of ROS induced oxidative damage.

If all those protective measures fail the organisms will ultimately encounter oxidative stress, which in turn is known to activate distinct genetic programs or the expression of certain genes (e.g. the general or oxidative stress response), to counteract the damaging effects by conversion and/or quenching of singlet oxygen and radical species [138-140].

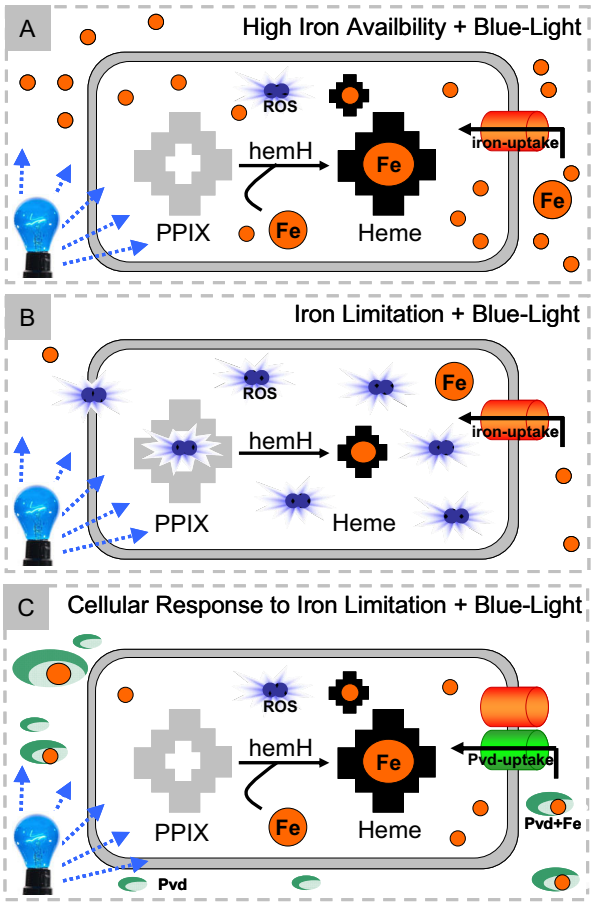


Figure 15: Hypothesis regarding the biological significance of blue-light under iron-limiting condition. A) If iron is sufficiently available to the bacterial cell iron-uptake to the cytoplasm will occur via the action of several receptors and transporters. Consequently, the heme biosynthesis pathway will be balanced and the production of reactive oxygen species (ROS) from endogenous photosensitizers (e.g. heme precursors, porphyrins) will be limited. B) Due to the lack of iron the heme biosynthesis pathway will be imbalanced at the final-step, namely the incorporation of ferrous-iron into the porphyrin-ring system of protoporphyrin IX (PPIX) by the action of a ferrochelatase. This will result in the accumulation of PPIX that can act as photosensitizer and eventually results in the formation ROS. C) In *P. putida KT2440,* an increased production of the iron-siderophore pyoverdine (pvd) is observed under blue-light illumination. Pvd very effectively scavenges even trace amounts of ferric-iron from the environment. Therefore, this cellular response might counteract the harmful effects of ROS formation be feeding iron into the bottleneck of the heme biosynthesis, namely the incorporation of iron into PPIX. Consequently, the heme biosynthesis pathway will be balanced toward the product (heme) and no or only limited accumulation of PPIX and thus ROS will occur.

According to the discussed hypothesis, the observed responses of *P. putida* KT2440 toward blue-light might all be part of cellular counteractive measures to avoid and to cope with oxidative damage in the cell due to PPIX or other endogenous

photosensitizer mediated ROS formation. An additional route, not illustrated in Figure 14, which might be involved in the control of photosensitizer-mediated ROS formation in *P. putida,* is depicted in Figure 15. This hypothesis is based on the observations described for the blue-light dependency of certain iron-limitation related cellular processes. Firstly, the observed increase in the pyoverdine production under blue-light might facilitate effective scavenging of even trace amounts of iron to feed into the bootleneck of heme biosynthesis, namely the incorporation of iron into the PPIX porphyrin ring system. And secondly, the increase in biofilm formation under blue-light illumination might be another way for the cells to overcome the direct exposure to blue-light. Accordingly, cells that are buried within the biofilm will obviously be less exposed to blue-light and moreover oxygen, and thus an increase in biofilm formation might increase the survivability of the strain under blue-light induced oxidative stress.

5.3 Application of bacterial LOV proteins

5.3.1 Reporter proteins for *in vivo* fluorescence without oxygen

Fluorescent reporter proteins such as the green fluorescent protein (GFP) originally identified from the jellyfish *Aequorea victoria*, its various variants that were developed in recent years, as well as newly identified GFP-related proteins from other marine species are valuable molecular tools in the non-invasive imaging of living specimens and cellular processes. Especially the availability of sophisticated color-variants for several fluorescent proteins that are emitting light from the blue to the red[141], the development of photo-switchable (on/off [142], red/green [143]) fluorescent markers, as well as variants of those proteins that show high fluorescent quantum yields and extended photostabilities [144-146], render them indispensable tools for the study of various molecular processes. For example- apart from simple labeling-purposes such as e.g. fluorescent-tagging of cells, tissues and microbial species - they can by employed for the study of gene expression, protein-localization, movement and interaction of proteins [141]. However, all members of the GFP family of fluorescent proteins possess one major drawback- they strictly depend on molecular oxygen as a co-factor for the synthesis of their light-emitting chromophores [147] and hence their use is restricted to biological systems that contain at least a certain amount of oxygen. This renders many interesting biological systems, such as microbial biofilms, strictly anaerobic microbes as well as oxygen deprived zones in cellular tissues difficult to analysze using GFP-related fluorescent reporter proteins. In Chapter 4.1 the development of fluorescent reporter proteins, that are based on bacterial LOV photoreceptor modules, is described. Based on their oxygen independent flavin photochemistry, those LOV-based fluorescent proteins should enable the study of molecular processes and living species under strict anaerobic conditions, and hence tackle the major drawback of GFP-related fluorescent reporters.

Application and optimization of bacterial LOV-based photosensor modules as fluorescent reporter systems for anaerobic systems

The LOV-based fluorescent reporter proteins described in Chapter 4.1 utilize the intrinsic autofluorescence of the flavin-chromophore that is bound within the photosensory LOV domain. Consequently, those proteins were termed FMN-based fluorescent proteins (FbFPs). In the wildtype LOV proteins, FMN that is in the dark noncovalently anchored in the LOV core, exclusively determines the LOV proteins

absorption characteristics. LOV protein bound FMN shows typical absorption maxima in the visible region at around 447nm and 475nm, attributed to the vibrational bands of the flavin $S_0 \rightarrow S_1$ transition [3, 148]. The absorption band in the UVA range ($S_0 \rightarrow S_2$ transition), is the most sensitive to the polarity of the microenvironment [149] and therefore differs slightly between different LOV proteins (e.g. in YtvA: 375nm [3], or in PpSB2-LOV: 360nm [65]). Therefore, the FMN molecule as light-absorbing chromophore can be excited by blue-light on a picosecond timescale to give rise to its triplet-state (LOV660) via Intersystem Crossing (ISC) from the corresponding singlet-state [62]. In phototropin LOVs about 50% of the excited singlet-state FMN molecules return spontaneously to the ground state [64] via radiative (fluorescence) and non-radiative (vibrational relaxation) processes. The radiative fraction of this decay process gives the LOV proteins their characteristic yellow-green fluorescence emission.

In the wildtype LOV domains of plant phototropin, the remaining 50% of the singlet-state FMN molecules enter the triplet-state via ISC and decay to give rise to the signaling state (LOV390) of the protein, which is characterized by the formation of a covalent bond between the FMN isoalloxazine ring C4a-atom and a conserved cysteine residue in the protein [64]. In the signaling state the FMN fluorescence is completely quenched and the sharp vibrationally resolved absorption band in the blue region is lost (see [69]). The signaling-state (LOV390) reverts in the dark via thermal processes slowly to the ground state. The slowest and therefore rate limiting step in the whole photocycle of the native LOV proteins is thus the dark recovery reaction. Thus, the relatively weak autofluorescence of the wildtype LOV proteins must be solely attributed to the radiative decay pathway of the FMN singlet-state directly to the ground-state, without concluding the photocycle via the LOV390 intermediate. Furthermore, with the non-fluorescent LOV390 intermediate as the longest living species in the LOV photocycle which only slowly reverts to the ground state, continuous excitation of the LOV protein will ultimately result in accumulation of the LOV390 intermediate and thus strongly inhibits the fluorescence of the wildtype LOV proteins. In order to generate a continuously autofluorescing marker protein, mutant derivatives of several bacterial LOV proteins were constructed in which the photoactive cysteine residue was mutationally replaced with a non-polar alanine. The general mutational strategy is depicted in Figure 16.

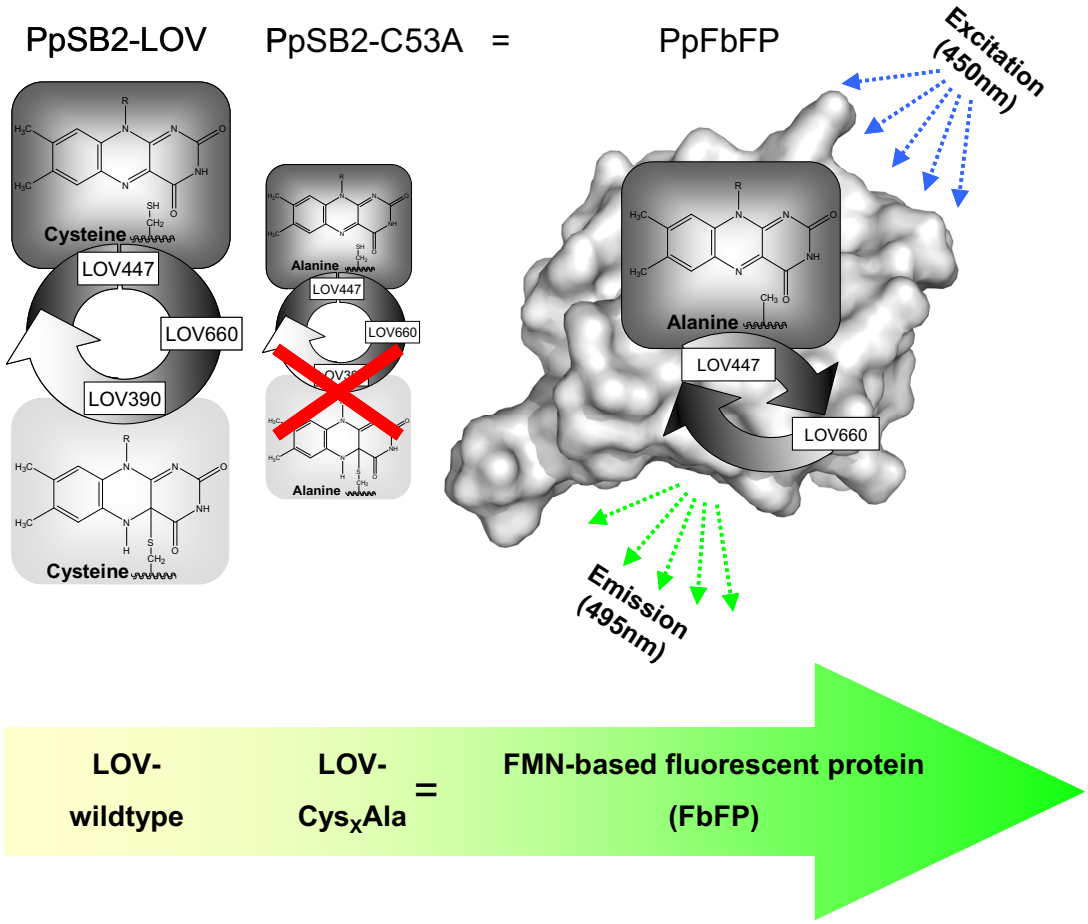


Figure 16: Mutational strategy for the generation of FMN-based fluorescent proteins on the basis of bacterial LOV photosensor module. Principle of the fluorescence emission from wildtype and mutated LOV proteins. The mutational replacement of the photoactive cysteine residue (e.g. Cys53 in PpSB2-LOV) results in an increased fluorescence quantum yield for the mutant LOV proteins due to reduced quenching of the chromophore (FMN) and concomitant inhibition of the LOV-photocycle. The fast switching between ground (LOV447) and FMN triplet state (LOV660) in the mutant proteins (=FbFPs) results in a continuous bright yellow-green fluorescence emission of the mutated LOV proteins as shown earlier.

Previous studies have demonstrated that this mutation in the LOV2 domain of phototropin phot1 from *Avena sativa* increases fluorescence emission, that results from both reduced fluorescence quenching of the FMN chromophore and inhibition of the LOV2 photocycle [64]. Both effects were thus utilized in order to generate continuously fluorescing FMN-based reporter proteins (FbFPs). Figure 17 exemplarily shows the fluorescence signal obtained from *Escherichia coli* cells, expressing one of the fluorescent reporter proteins described in Chapter 4.1.

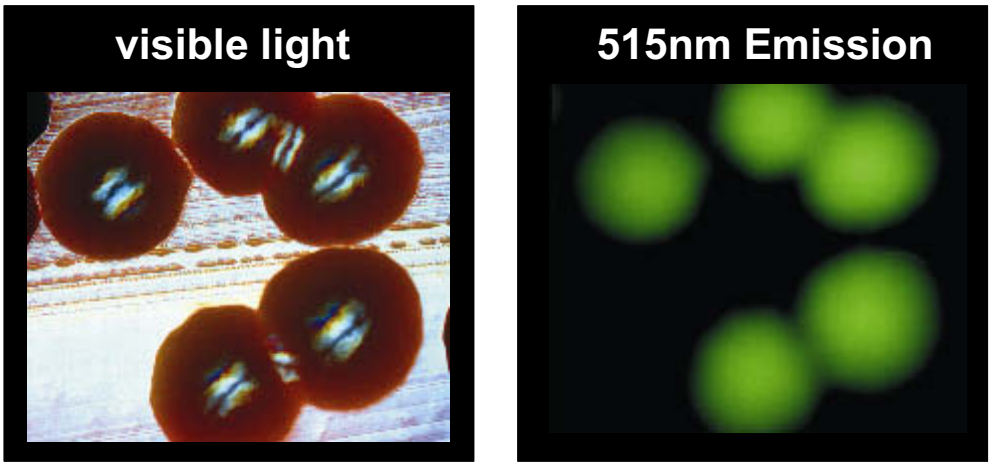


Figure 17: Fluorescing *E.coli* colonies. *E.coli* colonies expressing the FMN-based fluorescent protein (PpFbFP) grown on the solid-surface of an agar-plate. The figure shows two photographs of *E. coli* colonies expressing the novel fluorescence reporter protein PpFbFP from *P. putida*. On the left, bacterial colonies are depicted which were illuminated with full spectrum light. On the right hand side, the same colonies show PpFbFP-mediated green fluorescence (colonies were exited with blue light (488nm) and *in vivo* fluorescence was detected at a wavelength of 515nm).

LOV-based fluorescent reporter proteins can be used in anoxygenic systems without loss of the fluorescence signal

The facultative anaerobic phototroph bacterium *Rhodobacter capsulatus* B10S [150] was used as a "proof-of-principle" system to elucidate the applicability of the novel FMN-based fluorescence reporters. In particular, the fluorescent reporter protein, based on the PpSB2-LOV mutant PpSB2-LOVC53A (for simplicity named PpFbFP, whereas FbFP stands for FMN-based fluorescent protein) was further characterized regarding its suitability as a fluorescent reporter.

Whole cell (*in vivo*) fluorescence measurements were conducted under anaerobic condition, using a fluorimeter and additionally by employing confocal laser scanning microscopy in order to visualize the fluorescence of single bacterial cells. In both experimental settings a fluorescence signal under anaerobic conditions was only detectable from the FMN-based reporter system (PpFbFP), whereas for yellow fluorescent protein (YFP) containing cells no fluorescence could be monitored.

Furthermore, fluorescence quantum yields as well as molar extinction coefficients were determined for the different LOV-based fluorescent reporter proteins (FbFPs). Those analyses revealed a relative fluorescence brightness (defined as the fluorescence quantum yield multiplied by the molar extinction coefficient of the protein at the excitation wavelength) for the FbFPs that is within the same range as determined for the blue (BFP) and cyan (CFP) fluorescent variants of GFP, thus corroborating their suitability for non-invasive *in vivo* imaging under anaerobic

conditions. Since its first discovery in 1962 by Shimomura et al. [151] as a companion protein to aequorin, the famous chemiluminescent protein from the jellyfish *Aequorea victoria*, GFP has leaped from obscurity to become one of the most widely studied and exploited proteins in biochemistry and cell biology [147]. Its wide application in today's life-science sector has greatly benefited from the generation of various GFP mutants as elaborated above. Therefore, the initial characterization and "proof-of-principle" application of FbFPs as it is described in Chapter 4.1, should be followed by experiments to further optimize the proteins with regard to characteristics such as fluorescence brightness, photostability, and oligomerization state.

In conclusion, those novel FMN-based reporter systems (FbFPs) should hold a tremendous potential for both biotechnological and biomedical applications as well as for basic science as they will facilitate non-invasive imaging of living specimens and processes that occur under anaerobic conditions. For example the study of microbial biofilms, of strictly anaerobic microbes as well as the study of hypoxic regions within tissues and organs, should greatly benefit from the fluorescent reporter proteins described in Chapter 4.1.

5.3.2 Future Perspectives: Mutagenesis to probe bacterial LOV protein *in vivo* functionality and to optimize their applicability as fluorescent markers

Future research in the field of bacterial LOV photoreceptors might focus on several questions that remained elusive, despite the growing interest in this photoreceptor family during the last years.

i) How is the signal, that is initially received in the LOV domain, conveyed to fused or associated output domains e.g. in the blue-light photoreceptor YtvA of *B. subtilis*?

ii) How does the dark recovery reaction proceed mechanistically, and which amino acid residue(s) or protein regions determine the recovery rate constants that can differ significantly even between twin proteins that show a relatively high degree of sequence conservation, such as the PpSB1-LOV / PpSB2-LOV system of *P. putida*?

iii) Have the dark recovery and signal-transduction mechanisms been conserved throughout evolution between the plant phot-LOV systems and their bacterial ancestors?

Future Perspectives: Mutagenesis to probe the signal-relay and ligand binding in bacterial LOV photoreceptors

As discussed in the previous chapters, the answer to some of those pressing questions might be experimentally addressed using site-directed or random mutagenesis. Site-directed mutagenesis has been the workhorse for protein structure/function analysis over the past several decades [152]. The technique enables selective alteration of amino acid side-chains [152-154] e.g. with the purpose of assessing the energetics and structural determinants of protein-protein interactions [155] or protein-ligand interactions [156]. In particular, alanine scanning mutagenesis has been proven powerful for this purpose [157-159]. In an alanine-scanning mutagenesis approach, each and every single amino acid of the protein is consecutively substituted against an alanine residue that possesses only a methyl-group as side-chain. Therefore, the mutational replacement removes the side-chain

specific functionality from the replaced residue. In turn, the single mutations should allow conclusions regarding the functional role of specific amino acid residues in the protein (e.g. residues involved in ligand-binding). For example, in case of a cysteine to alanine mutation, the replacement will result in the exchange of the reactive thiol side-chain (in the cysteine) with the inert methyl group (in the alanine) and thus should abolish the cysteine-thiol specific functionality of the replaced residue.

Mutagenesis to probe the structure-function relationship in YtvA

As discussed in Chapter 1.6.1 YtvA functions in the environmental pathway of the σ^B controlled general response in *B. subtilis.* The STAS domain of YtvA is implied as the partner-switching protein domain that probably interacts via direct protein-protein interaction with the stressosome in a blue-light dependent manner eventually leading to σ^B activation [91].

But is this the sole role of YtvA in *B. subtilis*? Recent studies by Losi and co-workers suggested an additional functional role for YtvA and moreover for YtvAs STAS-domain [160]. It could be demonstrated experimentally that YtvA can bind a fluorescent derivative of guanosine triphosphate (GTP_{TR}) which can be replaced by both ATP and GTP *in vitro* [160]. Hence, providing an experimental proof for an NTP binding function of YtvA, which previously has been postulated solely based on the identification of two classical GTP-binding motifs DxxG and NKxD [161, 162] in the STAS domain, although the mode and position of the observed NTP binding in YtvA-STAS remains unclear. Notably, a GTP-binding function has also been demonstrated for the STAS protein SpoIIAA, that is involved in the regulation of the sporulation process in *B. subtilis* [163]. The authors consequently argued, that YtvA, in addition to its functional role in the general stress response pathway of *B. subtilis*, might be involved in the sporulation process, as both responses are related to a drastic drop of ATP [164] and GTP [165], respectively. In fact this dual role, if experimentally proved, might be attributed to similiar partner switching mechanisms between the STAS-domain of YtvA and the stressosome (to facilitate σ^B activation) and respectively between the STAS-domain and the intricate partner-switching network constituted by SpoIIAA, SpoIIAB and SpoIIE regulating σ^F activity [166] (to control compartment specific transcription during sporulation [166]). Henceforth, the alternative sigma-factors σ^B and σ^F might be controlled in part by proteins that are paralogous between the two pathways, thus providing a functional basis for a crosstalk between stress

response and sporulation. It is interesting to note that crosstalk between the two alternative pathways has in fact been observed under certain conditions[167].

In order to probe the involved signal-transduction mechanisms, structure-function dependency studies might be employed, e.g. sequence-scanning (e.g. alanine scanning) mutagenesis can be used to identify residues in both the LOV and STAS domain that are critical for the biological response or output (i.e. NTP-binding, σ^B activation etc.).

The first step in such a challenging mutagenesis project, that for example in case of YtvA (consisting of 261 aa) would involve the generation of 260 (excluding the start-codon) site-directed variants, is the design of a large number of mutagenic oligonucleotides to facilitate the mutagenesis of the *ytvA* gene by employing the Polymerase Chain Reaction (PCR). This design procedure can be cumbersome, laborious, and prone to error if performed by manually, but is easily amenable to automation without the necessity of expensive robotic equipment. In order to speed-up this process a computer tool was developed that enables the automatic design of oligonuclotides for such an sequence scanning/or saturation mutagenesis project. In Chapter 4.2 the corresponding computer tool (*insilico.mutagensis*) is described. The tool is written entirely in Perl and implements a MySQL database for easy data-storage and processing. Moreover, a user-friendly HTML-interface, accessible via the world wide web (www.insilico.uni-duesseldorf.de) facilitates system-independent usability. Moreover, every directed evolution project involving sequence saturation or the directed scanning of protein regions, e.g. to probe receptor affinity, ligand binding or catalytic functions, should benefit in principal from an automatic oligonucleotide design procedure as it will speed up the whole process from PCR-driven mutagenesis to the identification of protein variants exhibiting the desired catalytic function or *in vitro* biological effect (e.g. receptor / ligand binding).

Alanine scanning of the GTP-binding motif of YtvA

A large scale, alanine-scanning mutagenesis project was initiated to completely scan the YtvA protein sequence, by consecutively replacing each and every single amino acid in the protein with a non-polar alanine residue, thereby removing the side-chain specific functionality from the replaced residue as explained above.

As a first test system, experiments using four selected alanine scanning mutants of the $D_{191}SLG_{194}$-GTP binding motif in YtvA, were conducted to probe the binding-mode of the fluorescent GTP-analogue (GTP-TR) in an experimental setting as described by Buttani *et al.* [160]. For this purpose, mutagenic oligonucleotides were computationally generated using the above described *insilico.mutagensis* tool, which was developed and is presented as part of this thesis (see Chapter 4.2).

GTP-binding experiments were performed in cooperation with the group of Dr. Losi (University of Parma, Italy) and Prof. Gaertner (MPI Bioinorganic Chemistry, Muelheim a.d.Ruhr, Germany).

In those, yet unpublished studies, the $D_{191}SLG_{194}$-motif of YtvA was mutationally scanned, meaning that the four amino acid positions in the sequence-motif were consecutively substituted against a non-polar alanine residue. The corresponding purified mutant proteins were analyzed for their ability to bind GTP-TR, similarly to the experiments described in [160]. Notably, the mutation S193A apparently impairs GTP-TR binding to YtvA, without affecting protein-folding and photocycle integrity. This provides experimental evidence that the GTP-binding motif (DSLG), which was previously identified using bioinformatics [3] is involved in the GTP-binding to YtvA. As part of this large scale mutagenesis project, at present 130 mutants could be generated. Those variants cover the complete STAS domain of YtvA. In order to completely scan the YtvA protein, the missing approximately 130 mutations that cover the LOV domain and the helical linker portion of YtvA have still to be generated. All those mutant proteins, if completely generated and screened, should enable us to narrow down the position of GTP-binding in the STAS domain of YtvA and furthermore, would enable us to elucidate the role of YtvAs LOV domain in the process.

Furthermore, with the advent of functional assays to probe the blue-light dependent physiological role of YtvA that were published during the last year [90, 91], it should become possible to conduct structure-function studies *in vivo*, e.g. by using site-directed mutagenesis to elucidate the role of YtvAs STAS domain, as well as to probe the signal-relay between LOV and STAS, in the biological response.

Random mutagenesis to probe the mechanistics of the LOV dark recovery reaction

Random mutagenesis of, e.g. PpSB1-LOV and subsequent screening of the variants on a fast dark recovery reaction, should provide further insight into the mechanistics of the dark recovery process by identifying hot-spot residues that control the (catalytic) back reaction. In a similar directed evolution approach Christie and co-workers, recently identified hot-spot residues involved in the dark recovery of the Asphot1-LOV2 domain [108]. This is discussed in more detail in Chapter 3.3 and Chapter 5.2.3. In brief, the protein region that was identified in the plant LOV domain as hot-spot controlling the dark recovery reaction is conserved (identical) between the twin bacterial LOV proteins (PpSB1-LOV and PpSB2-LOV). With regard to the observation that those twin LOV proteins nevertheless show dramatically different recovery kinetics, it seems unlikely that the same mechanisms as discussed by Christie et al. [108] are involved in the latter mentioned bacterial LOV systems. Notwithstanding the fact that similar experiments were already conducted for a plant LOV domain, random mutagenesis might all the same provide a fast and easy approach to study the underlying principles of the dark recovery reaction in bacterial LOV systems. Those experiments might thus enable us in the future to shed some light on the conservation of the mechanisms that govern the LOV domain dark recovery in plant and bacterial photoreceptor systems alike.

Future Perspectives: The use of mutagenesis to optimize LOV-based fluorescent reporters

Furthermore, with respect to the biotechnological application of bacterial LOV proteins as fluorescent reporter proteins in anaerobic systems as described in Chapters 4.1 and 5.3.1, site-directed mutagenesis (e.g. sequence scanning/saturation) as well as random mutagenesis might be employed to further optimize certain protein properties (e.g. oligomerization) and fluorescence characteristics (fluorescence brightness and photostability) of the respective fluorescent reporter, in a similar fashion as it has been done for GFP and its various mutants. Those experiments might, in the future, greatly benefit from an automated design procedure for the necessary mutagenic PCR-oligonuclueotides, as described in Chapter 4.2.

For example experiments are conceivable that would mutationally scan the dimer-interface of the isolated LOV domain of BsFbFP or PpFbFP, which both form stable homo-dimers in solution (as discussed in Chapter 4.1), with the goal to disrupt the respective fluorescent reporter- protein oligomers.

It is generally accepted knowledge that fluorescence reporters with monomeric tertiary structure are preferable over dimeric or even tetrameric organizations.

For example native DsRed (a red-fluorescent reporter), was shown to be tetrameric in its matured fluorescent state [168]. This assembly, however, causes everything to which DsRed is fused to aggregate, which is disastrous for many experimental settings [169]. The problem could however be overcome by using a directed evolution approach which allowed the generation of a monomeric form of DsRed, by interrupting oligomer contacts [170]. Comparable optimization strategies would be feasible for BsFbFP or PpFbFp, with the aim to generate photostable, highly-fluorescent, reporter proteins that are monomeric in their tertiary structure. Consequently, those should in principal be more easily applicable as fluorescent translational-fusion reporters.

6 References

1. **Christie, J. M., P. Reymond, G. K. Powell, P. Bernasconi, A. A. Raibekas, E. Liscum, and W. R. Briggs.** 1998. *Arabidopsis* NPH1: a flavoprotein with the properties of a photoreceptor for phototropism. Science **282:**1698-701.

2. **Christie, J. M.** 2007. Phototropin blue-light receptors. Annu Rev Plant Biol **58:**21-45.

3. **Losi, A., E. Polverini, B. Quest, and W. Gärtner.** 2002. First evidence for phototropin-related blue-light receptors in prokaryotes. Biophys J **82:**2627-34.

4. **Losi, A.** 2006. Flavin-based photoreceptors in bacteria *In* D. Haeder and G. Jori (ed.), Flavin photochemistry and photobiology. Elsevier, Amsterdam.

5. **Matsuoka, D., T. Iwata, K. Zikihara, H. Kandori, and S. Tokutomi.** 2006. Primary Processes During the Light-signal Transduction of Phototropin. Photochem Photobiol. **83:**122-33.

6. **Crosson, S., and K. Moffat.** 2001. Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. Proc Natl Acad Sci U S A **98:**2995-3000.

7. **Nakasako, M., T. Iwata, D. Matsuoka, and S. Tokutomi.** 2004. Light-induced structural changes of LOV domain-containing polypeptides from *Arabidopsis* phototropin 1 and 2 studied by small-angle X-ray scattering. Biochemistry **43:**14881-90.

8. **Nakasako, M., D. Matsuoka, K. Zikihara, and S. Tokutomi.** 2005. Quaternary structure of LOV-domain containing polypeptide of Arabidopsis FKF1 protein. FEBS Lett **579:**1067-71.

9. **Hall D.D.,** 1999. Photosynthesis *In* K. K. Rao (hrsg), Studies in Biology, Cambridge University Press, Cambridge

10. 2006. ASTM G173-03e1, "Standard Tables for Reference Solar Spectral Irradiances: Direct Normal and Hemispherical on 37° Tilted Surface, ASTM International. Book of Standards vol. 14.04. ASTM International, West Conshohocken, PA, USA.

11. 1976. U.S. Standard Atmosphere. United States Committee on Extension to the Standard Atmosphere, National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, United States Air Force, Washington D.C.

12. **Wald, G.** 1968. The molecular basis of visual excitation. Nature **219:**800-7.

13. **Spudich, J. L., C. S. Yang, K. H. Jung, and E. N. Spudich.** 2000. Retinylidene proteins: structures and functions from archaea to humans. Annu Rev Cell Dev Biol **16:**365-92.

14. **Quail, P. H.** 1998. The phytochrome family: dissection of functional roles and signalling pathways among family members. Philos Trans R Soc Lond B Biol Sci **353:**1399-403.

15. **Kort, R., W. D. Hoff, M. Van West, A. R. Kroon, S. M. Hoffer, K. H. Vlieg, W. Crielaand, J. J. Van Beeumen, and K. J. Hellingwerf.** 1996. The xanthopsins: a new family of eubacterial blue-light photoreceptors. Embo J **15:**3209-18.

16. **Huala, E., P. W. Oeller, E. Liscum, I. S. Han, E. Larsen, and W. R. Briggs.** 1997. Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain. Science **278:**2120-3.

17. **Somers, D. E., T. F. Schultz, M. Milnamow, and S. A. Kay.** 2000. ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. Cell **101:**319-29.

18. **Ahmad, M., and A. R. Cashmore.** 1993. HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. Nature **366:**162-6.

19. **Gomelsky, M., and G. Klug.** 2002. BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms. Trends Biochem Sci **27:**497-500.

20. **Christie, J. M., M. Salomon, K. Nozue, M. Wada, and W. R. Briggs.** 1999. LOV (light, oxygen, or voltage) domains of the blue-light photoreceptor phototropin (nph1): binding sites for the chromophore flavin mononucleotide. Proc Natl Acad Sci U S A **96:**8779-83.

21. **Oesterhelt, D.** 1998. The structure and mechanism of the family of retinal proteins from halophilic archaea. Curr Opin Struct Biol **8:**489-500.

22. **Schegk, E. S., and D. Oesterhelt.** 1988. Isolation of a prokaryotic photoreceptor: sensory rhodopsin from halobacteria. Embo J **7:**2925-33.

23. **Stoeckenius, W.** 1985. The rhodopsin-like pigments of halobacteria: light-energy and signal transducers in an archaebacterium. Trends Biochem Sci **10:**483-6.

24. **Sineshchekov, O. A., K. H. Jung, and J. L. Spudich.** 2002. Two rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci U S A **99:**8689-94.

25. **Yoshizawa, T.** 1984. Photophysiological functions of visual pigments. Adv Biophys **17:**5-67.

26. **He, S., W. Dong, Q. Deng, S. Weng, and W. Sun.** 2003. Seeing more clearly: recent advances in understanding retinal circuitry. Science **302:**408-11.

27. **Barinaga, M.** 2002. Circadian clock. How the brain's clock gets daily enlightenment. Science **295:**955-7.

28. **Ruiz-Gonzalez, M. X., and I. Marin.** 2004. New insights into the evolutionary history of type 1 rhodopsins. J Mol Evol **58:**348-58.

29. **Jung, K. H., V. D. Trivedi, and J. L. Spudich.** 2003. Demonstration of a sensory rhodopsin in eubacteria. Mol Microbiol **47:**1513-22.

30. **Purschwitz, J., S. Muller, C. Kastner, and R. Fischer.** 2006. Seeing the rainbow: light sensing in fungi. Curr Opin Microbiol **9:**566-71.

31. **Karniol, B., J. R. Wagner, J. M. Walker, and R. D. Vierstra.** 2005. Phylogenetic analysis of the phytochrome superfamily reveals distinct microbial subfamilies of photoreceptors. Biochem J **392:**103-16.

32. **Rockwell, N. C., and J. C. Lagarias.** 2006. The structure of phytochrome: a picture is worth a thousand spectra. Plant Cell **18:**4-14.

33. **Mathews, S.** 2006. Phytochrome-mediated development in land plants: red light sensing evolves to meet the challenges of changing light environments. Mol Ecol **15:**3483-503.

34. **Montgomery, B. L.** 2007. Sensing the light: photoreceptive systems and signal transduction in cyanobacteria. Mol Microbiol **64:**16-27.

35. **Kehoe, D. M., and A. R. Grossman.** 1996. Similarity of a chromatic adaptation sensor to phytochrome and ethylene receptors. Science **273:**1409-12.

36. **Wilde, A., Y. Churin, H. Schubert, and T. Borner.** 1997. Disruption of a *Synechocystis* sp. PCC 6803 gene with partial similarity to phytochrome genes alters growth under changing light qualities. FEBS Lett **406:**89-92.

37. **Meyer, T. E.** 1985. Isolation and characterization of soluble cytochromes, ferredoxins and other chromophoric proteins from the halophilic phototrophic bacterium *Ectothiorhodospira halophila*. Biochim Biophys Acta **806:**175-83.

38. **Kyndt, J. A., T. E. Meyer, and M. A. Cusanovich.** 2004. Photoactive yellow protein, bacteriophytochrome, and sensory rhodopsin in purple phototrophic bacteria. Photochem Photobiol Sci **3:**519-30.

39. **Sprenger, W. W., W. D. Hoff, J. P. Armitage, and K. J. Hellingwerf.** 1993. The eubacterium *Ectothiorhodospira halophila* is negatively phototactic, with a wavelength dependence that fits the absorption spectrum of the photoactive yellow protein. J Bacteriol **175:**3096-104.

40. **Kirk, J. T. O.** 1994. The nature of the underwater light field, p. 129-168, Light and photosynthesis in aquatic ecosystems, 2nd ed. Cambridge University Press, Cambridge

41. **Pattison, D. I., and M. J. Davies.** 2006. Actions of ultraviolet light on cellular structures. Exs**:**131-57.

42. **Ryter, S. W., H. P. Kim, A. Hoetzel, J. W. Park, K. Nakahira, X. Wang, and A. M. Choi.** 2007. Mechanisms of cell death in oxidative stress. Antioxid Redox Signal **9:**49-89.

43. **Spikes, J. D.** 1989. Photosensitization., p. 79-110. *In* K. C. Smith (ed.), The Science of Photobiology. Plenum Press, New York and London.

44. **Lin, C., and T. Todo.** 2005. The cryptochromes. Genome Biol **6:**220.

45. **Cashmore, A. R.** 2003. Cryptochromes: enabling plants and animals to determine circadian time. Cell **114:**537-43.

46. **Lin, C., and D. Shalitin.** 2003. Cryptochrome structure and signal transduction. Annu Rev Plant Biol **54:**469-96.

47. **Sancar, A.** 2003. Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. Chem Rev **103:**2203-37.

48. **van der Horst, M. A., and K. J. Hellingwerf.** 2004. Photoreceptor proteins, "star actors of modern times": a review of the functional dynamics in the structure of representative members of six different photoreceptor families. Acc Chem Res **37:**13-20.

49. **Hitomi, K., K. Okamoto, H. Daiyasu, H. Miyashita, S. Iwai, H. Toh, M. Ishiura, and T. Todo.**
2000. Bacterial cryptochrome and photolyase: characterization of two photolyase-like genes of *Synechocystis* sp. PCC6803. Nucleic Acids Res **28:**2353-62.

50. **Worthington, E. N., I. H. Kavakli, G. Berrocal-Tito, B. E. Bondo, and A. Sancar.** 2003. Purification and characterization of three members of the photolyase/cryptochrome family blue-light photoreceptors from *Vibrio cholerae*. J Biol Chem **278:**39143-54.

51. **Brudler, R., K. Hitomi, H. Daiyasu, H. Toh, K. Kucho, M. Ishiura, M. Kanehisa, V. A. Roberts, T. Todo, J. A. Tainer, and E. D. Getzoff.**
2003. Identification of a new cryptochrome class. Structure, function, and evolution. Mol Cell **11:**59-67.

52. **Green, C. B.** 2004. Cryptochromes: tail-ored for distinct functions. Curr Biol **14:** 847-9.

53. **Selby, C. P., and A. Sancar.** 2006. A cryptochrome/photolyase class of enzymes with single-stranded DNA-specific photolyase activity. Proc Natl Acad Sci U S A **103:**17696-700.

54. **Braatsch, S., M. Gomelsky, S. Kuphal, and G. Klug.** 2002. A single flavoprotein, AppA, integrates both redox and light signals in *Rhodobacter sphaeroides*. Mol Microbiol **45:**827-36.

55. **Iseki, M., S. Matsunaga, A. Murakami, K. Ohno, K. Shiga, K. Yoshida, M. Sugai, T. Takahashi, T. Hori, and M. Watanabe.** 2002. A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. Nature **415:**1047-51.

56. **Ntefidou, M., M. Iseki, M. Watanabe, M. Lebert, and D. P. Häder.** 2003. Photoactivated adenylyl cyclase controls phototaxis in the flagellate *Euglena gracilis*. Plant Physiol **133:**1517-21.

57. **Darwin, C. R. A. b. F. D.** 1880. *The power of movement in plants*. John Murray, London.

58. **Briggs, W. R., and J. M. Christie.** 2002. Phototropins 1 and 2: versatile plant blue-light receptors. Trends Plant Sci **7:**204-10.

59. **Ohgishi, M., K. Saji, K. Okada, and T. Sakai.** 2004. Functional analysis of each blue light receptor, cry1, cry2, phot1, and phot2, by using combinatorial multiple mutants in *Arabidopsis*. Proc Natl Acad Sci U S A **101:**2223-8.

60. **Huang, K., and C. F. Beck.** 2003. Phototropin is the blue-light receptor that controls multiple steps in the sexual life cycle of the green alga *Chlamydomonas reinhardtii*. Proc Natl Acad Sci U S A **100:**6269-74.

61. **Imaizumi, T., H. G. Tran, T. E. Swartz, W. R. Briggs, and S. A. Kay.** 2003. FKF1 is essential for photoperiodic-specific light signalling in *Arabidopsis*. Nature **426:**302-6.

62. **Kennis, J. T., S. Crosson, M. Gauden, I. H. van Stokkum, K. Moffat, and R. van Grondelle.** 2003. Primary reactions of the LOV2 domain of phototropin, a plant blue-light photoreceptor. Biochemistry **42:**3385-92.

63. **Kottke, T., J. Heberle, D. Hehn, B. Dick, and P. Hegemann.** 2003. Phot-LOV1: photocycle of a blue-light receptor domain from the green alga *Chlamydomonas reinhardtii*. Biophys J **84:**1192-201.

64. **Swartz, T. E., S. B. Corchnoy, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs, and R. A. Bogomolni.** 2001. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. J Biol Chem **276:**36493-500.

65. **Krauss, U., A. Losi, W. Gärtner, K.- E. Jaeger, and T. Eggert.** 2005. Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct. Phys Chem Chem Phys **7:**2804-11.

66. **Losi, A.** 2004. The bacterial counterparts of plants phototropins,. Photochem Photobiol Sci **3:**566-574.

67. **Zikihara, K., T. Iwata, D. Matsuoka, H. Kandori, T. Todo, and S. Tokutomi.** 2006. Photoreaction cycle of the light, oxygen, and voltage domain in FKF1 determined by low-temperature absorption spectroscopy. Biochemistry **45:**10828-37.

68. **Zoltowski, B. D., C. Schwerdtfeger, J. Widom, J. J. Loros, A. M. Bilwes, J. C. Dunlap, and B. R. Crane.** 2007. Conformational switching in the fungal light sensor Vivid. Science **316:**1054-7.

69. **Losi, A.** 2007. Flavin-based Blue-light Photosensors: A Photobiophysics Update Photochem Photobiol **doi: 10.1111/j.1751-1097.2007.00195.x**.

70. **Crosson, S., and K. Moffat.** 2002. Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch. Plant Cell **14:**1067-75.

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

72. **Freddolino, P. L., M. Dittrich, and K. Schulten.** 2006. Dynamic switching mechanisms in LOV1 and LOV2 domains of plant phototropins. Biophys J **91:**3630-9.

73. **Christie, J. M., T. E. Swartz, R. A. Bogomolni, and W. R. Briggs.** 2002. Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function. Plant J **32:**205-19.

74. **Jones, M. A., K. A. Feeney, S. M. Kelly, and J. M. Christie.** 2007. Mutational analysis of phototropin 1 provides insights into the mechanism underlying LOV2 signal transmission. J Biol Chem **282:**6405-14.

75. **Nozaki, D., T. Iwata, T. Ishikawa, T. Todo, S. Tokutomi, and H. Kandori.** 2004. Role of Gln1029 in the photoactivation processes of the LOV2 domain in *Adiantum* phytochrome3. Biochemistry **43:**8373-9.

76. **Harper, S. M., L. C. Neil, and K. H. Gardner.** 2003. Structural basis of a phototropin light switch. Science **301:**1541-4.

77. **Harper, S. M., J. M. Christie, and K. H. Gardner.** 2004. Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity. Biochemistry **43:**16184-92.

78. **Hughes, J., T. Lamparter, F. Mittmann, E. Hartmann, W. Gärtner, A. Wilde, and T. Borner.** 1997. A prokaryotic phytochrome. Nature **386:**663.

79. **Gärtner, W., and A. Losi.** 2003. Crossing the borders: archaeal rhodopsins go bacterial. Trends Microbiol **11:**405-7.

80. **Montgomery, B. L., and J. C. Lagarias.** 2002. Phytochrome ancestry: sensors of bilins and light. Trends Plant Sci **7:**357-66.

81. **Cheng, P., Q. He, Y. Yang, L. Wang, and Y. Liu.** 2003. Functional conservation of light, oxygen, or voltage domains in light sensing. Proc Natl Acad Sci U S A **100:**5938-43.

82. **Cao, Z., V. Buttani, A. Losi, and W. Gärtner.** 2007. A blue light inducible two component signal transduction system in the plant pathogen *Pseudomonas syringae* pv. tomato. Biophys J.

83. **Swartz, T. E., T. S. Tseng, M. A. Frederickson, G. Paris, D. J. Comerci, G. Rajashekara, J. G. Kim, M. B. Mudgett, G. A. Splitter, R. A. Ugalde, F. A. Goldbaum, W. R. Briggs, and R. A. Bogomolni.** 2007. Blue-light-activated histidine kinases: two-component sensors in bacteria. Science **317:**1090-3.

84. **Chan, C., R. Paul, D. Samoray, N. C. Amiot, B. Giese, U. Jenal, and T. Schirmer.** 2004. Structural basis of activity and allosteric control of diguanylate cyclase. Proc Natl Acad Sci U S A **101:**17084-9.

85. **Galperin, M. Y., A. N. Nikolskaya, and E. V. Koonin.** 2001. Novel domains of the prokaryotic two-component signal transduction systems. FEMS Microbiol Lett **203:**11-21.

86. **Romling, U., M. Gomelsky, and M. Y. Galperin.** 2005. C-di-GMP: the dawning of a novel bacterial signalling system. Mol Microbiol **57:**629-39.

87. **Aravind, L., V. Anantharaman, S. Balaji, M. M. Babu, and L. M. Iyer.** 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol Rev **29:**231-62.

88. **Hefti, M. H., K. J. Francoijs, S. C. de Vries, R. Dixon, and J. Vervoort.** 2004. The PAS fold. A redefinition of the PAS domain based upon structural prediction. Eur J Biochem **271:**1198-208.

89. **Aravind, L., and C. P. Ponting.** 1997. The GAF domain: an evolutionary link between diverse phototransducing proteins. Trends Biochem Sci **22:**458-9.

90. **Avila-Perez, M., K. J. Hellingwerf, and R. Kort.** 2006. Blue light activates the sigmaB-dependent stress response of *Bacillus subtilis* via YtvA. J Bacteriol **188:**6411-4.

91. **Gaidenko, T. A., T. J. Kim, A. L. Weigel, M. S. Brody, and C. W. Price.** 2006. The blue-light receptor YtvA acts in the environmental stress signaling pathway of *Bacillus subtilis*. J Bacteriol **188:**6387-95.

92. **Suzuki, N., N. Takaya, T. Hoshino, and A. Nakamura.** 2007. Enhancement of a sigma(B)-dependent stress response in *Bacillus subtilis* by light via YtvA photoreceptor. J Gen Appl Microbiol **53:**81-8.

93. **Mizoguchi, T., and G. Coupland.** 2000. ZEITLUPE and FKF1: novel connections between flowering time and circadian clock control. Trends Plant Sci **5:**409-11.

94. **Ballario, P., P. Vittorioso, A. Magrelli, C. Talora, A. Cabibbo, and G. Macino.** 1996. White collar-1, a central regulator of blue light responses in *Neurospora*, is a zinc finger protein. Embo J **15:**1650-7.

95. **Talora, C., L. Franchi, H. Linden, P. Ballario, and G. Macino.** 1999. Role of a white collar-1-white collar-2 complex in blue-light signal transduction. Embo J **18:**4961-8.

96. **Hecker, M., J. Pane-Farre, and U. Volker.** 2006. SigB-Dependent General Stress Response in *Bacillus subtilis* and Related Gram-Positive Bacteria. Annu Rev Microbiol.

97. **Haldenwang, W. G.** 1995. The sigma factors of *Bacillus subtilis*. Microbiol Rev **59:**1-30.

98. **Akbar, S., T. A. Gaidenko, C. M. Kang, M. O'Reilly, K. M. Devine, and C. W. Price.** 2001. New family of regulators in the environmental signaling pathway which activates the general stress transcription factor sigma(B) of *Bacillus subtilis*. J Bacteriol **183:**1329-38.

99. **Reeves, A., and W. G. Haldenwang.** 2007. Isolation and characterization of dominant mutations in the *Bacillus subtilis* stressosome components RsbR and RsbS. J Bacteriol **189:**1531-41.

100. **Drepper, T., T. Eggert, F. Circolone, A. Heck, U. Krauss, J. K. Guterl, M. Wendorff, A. Losi, W. Gärtner, and K.-E. Jaeger.** 2007. Reporter proteins for in vivo fluorescence without oxygen. Nat Biotechnol **25:**443-5.

101. **Bryant, D. A., and N. U. Frigaard.** 2006. Prokaryotic photosynthesis and phototrophy illuminated. Trends Microbiol **14:**488-96.

102. **Losi, A., B. Quest, and W. Gärtner.** 2003. Listening to the blue: the time-resolved thermodynamics of the bacterial blue-light receptor YtvA and its isolated LOV domain. Photochem Photobiol Sci **2:**759-66.

103. **Kim, W. Y., S. Fujiwara, S. S. Suh, J. Kim, Y. Kim, L. Han, K. David, J. Putterill, H. G. Nam, and D. E. Somers.** 2007. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. Nature **449:**356-60.

104. **Sawa, M., D. A. Nusinow, S. A. Kay, and T. Imaizumi.** 2007. FKF1 and GIGANTEA Complex Formation is Required for Day-Length Measurement in *Arabidopsis*. Science.

105. **Buttani, V., A. Losi, T. Eggert, U. Krauss, K.- E. Jaeger, Z. Cao, and W. Gärtner.** 2007. Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions. Photochem Photobiol Sci **6:**41-9.

106. **Möglich, A., and K. Moffat.** 2007. Structural Basis for Light-dependent Signaling in the Dimeric LOV Domain of the Photosensor YtvA. J Mol Biol **373:**112-26.

107. **Alexandre, M. T., J. C. Arents, R. van Grondelle, K. J. Hellingwerf, and J. T. Kennis.** 2007. A base-catalyzed mechanism for dark state recovery in the Avena sativa

phototropin-1 LOV2 domain. Biochemistry **46:**3129-37.

108. **Christie, J. M., S. B. Corchnoy, T. E. Swartz, M. Hokenson, I. S. Han, W. R. Briggs, and R. A. Bogomolni.** 2007. Steric Interactions Stabilize the Signaling State of the LOV2 Domain of Phototropin 1. Biochemistry.

109. **Losi, A., E. Ternelli, and W. Gärtner.** 2004. Tryptophan fluorescence in the *Bacillus subtilis* phototropin-related protein YtvA as a marker of interdomain interaction. Photochem Photobiol **80:**150-3.

110. **Kortemme, T., D. E. Kim, and D. Baker.** 2004. Computational alanine scanning of protein-protein interfaces. Sci STKE **2004:** l2.

111. **Eitoku, T., Y. Nakasone, D. Matsuoka, S. Tokutomi, and M. Terazima.** 2005. Conformational dynamics of phototropin 2 LOV2 domain with the linker upon photoexcitation. J Am Chem Soc **127:**13238-44.

112. **Nakasone, Y., T. Eitoku, D. Matsuoka, S. Tokutomi, and M. Terazima.** 2007. Dynamics of conformational changes of *Arabidopsis* phototropin 1 LOV2 with the linker domain. J Mol Biol **367:**432-42.

113. **Eitoku, T., Y. Nakasone, K. Zikihara, D. Matsuoka, S. Tokutomi, and M. Terazima.** 2007. Photochemical intermediates of *Arabidopsis* phototropin 2 LOV domains associated with conformational changes. J Mol Biol **371:**1290-303.

114. **Matsuoka, D., and S. Tokutomi.** 2005. Blue light-regulated molecular switch of Ser/Thr kinase in phototropin. Proc Natl Acad Sci U S A **102:**13337-42.

115. **Bayles, K. W.** 2007. The biological role of death and lysis in biofilm development. Nat Rev Microbiol **5:**721-6.

116. **Rice, K. C., E. E. Mann, J. L. Endres, E. C. Weiss, J. E. Cassat, M. S. Smeltzer, and K. W. Bayles.** 2007. The cidA murein hydrolase regulator contributes to DNA release and biofilm development in *Staphylococcus aureus*. Proc Natl Acad Sci U S A **104:**8113-8.

117. **Visca, P., F. Imperi, and I. L. Lamont.** 2007. Pyoverdine siderophores: from biogenesis to biosignificance. Trends Microbiol **15:**22-30.

118. **Meyer, J. M., A. Stintzi, D. De Vos, P. Cornelis, R. Tappe, K. Taraz, and H. Budzikiewicz.** 1997. Use of siderophores to type pseudomonads: the three *Pseudomonas aeruginosa* pyoverdine systems. Microbiology **143 (Pt 1):**35-43.

119. **Boukhalfa, H., S. D. Reilly, R. Michalczyk, S. Iyer, and M. P. Neu.** 2006. Iron (III) coordination properties of a pyoverdin siderophore produced by *Pseudomonas putida* ATCC 33015. Inorg Chem **45:**5607-16.

120. **Cody, Y. S., and D. C. Gross.** 1987. Characterization of Pyoverdin (pss), the Fluorescent Siderophore Produced by *Pseudomonas syringae* pv. syringae. Appl Environ Microbiol **53:**928-934.

121. **Poole, K., and G. A. McKay.** 2003. Iron acquisition and its control in *Pseudomonas aeruginosa*: many roads lead to Rome. Front Biosci **8:**d661-86.

122. **Tappe, R., Taraz, K., Budzikiewicz, H., Meyer, J.-M., Lefevre, J.-F., .** 1993. Structure elucidation of a pyoverdin produced by *Pseudomonas aeruginosa* ATCC 27853. J Prakt Chem **335:**83-87.

123. **Beare, P. A., R. J. For, L. W. Martin, and I. L. Lamont.** 2003. Siderophore-mediated cell signalling in *Pseudomonas aeruginosa*: divergent pathways regulate virulence factor production and siderophore receptor synthesis. Mol Microbiol **47:**195-207.

124. **Buyer, J. S., and J. Leong.** 1986. Iron transport-mediated antagonism between plant growth-promoting and plant-deleterious *Pseudomonas* strains. J Biol Chem **261:**791-4.

125. **Schroth, M. N., and J. G. Hancock.** 1982. Disease-Suppressive Soil and Root-Colonizing Bacteria. Science **216:**1376-1381.

126. **Jones, A. M., S. E. Lindow, and M. C. Wildermuth.** 2007. Salicylic acid, yersiniabactin, and pyoverdin production by the model phytopathogen *Pseudomonas syringae* pv. tomato DC3000: synthesis, regulation, and impact on tomato and Arabidopsis host plants. J Bacteriol **189:**6773-86.

127. **O'Toole, G. A., and R. Kolter.** 1998. Initiation of biofilm formation in *Pseudomonas fluorescens* WCS365 proceeds via multiple, convergent signalling pathways: a genetic analysis. Mol Microbiol **28:**449-61.

128. **Panek, H., and M. R. O'Brian.** 2002. A whole genome view of prokaryotic haem biosynthesis. Microbiology **148:**2273-82.

129. **Pourzand, C., O. Reelfs, E. Kvam, and R. M. Tyrrell.** 1999. The iron regulatory protein can determine the effectiveness of 5-aminolevulinic acid in inducing protoporphyrin IX in human primary skin fibroblasts. J Invest Dermatol **112:**419-25.

130. **Miyamoto, K., K. Nakahigashi, K. Nishimura, and H. Inokuchi.** 1991. Isolation and characterization of visible light-sensitive mutants of *Escherichia coli* K12. J Mol Biol **219:**393-8.

131. **Baysse, C., S. Matthijs, M. Schobert, G. Layer, D. Jahn, and P. Cornelis.** 2003. Co-ordination of iron acquisition, iron porphyrin chelation and iron-protoporphyrin export via the cytochrome c biogenesis protein CcmC in *Pseudomonas fluorescens*. Microbiology **149:**3543-52.

132. **Fritsch, C., G. Goerz, and T. Ruzicka.** 1998. Photodynamic therapy in dermatology. Arch Dermatol **134:**207-14.

133. **Peng, Q., T. Warloe, K. Berg, J. Moan, M. Kongshaug, K. E. Giercksky, and J. M. Nesland.** 1997. 5-Aminolevulinic acid-based photodynamic therapy. Clinical research and future challenges. Cancer **79:**2282-308.

134. **Armitage, J. P., and K. J. Hellingwerf.** 2003. Light-induced behavioral responses (phototaxis') in prokaryotes. Photosynth Res **76:**145-55.

135. **Takano, H., D. Asker, T. Beppu, and K. Ueda.** 2006. Genetic control for light-induced carotenoid production in non-phototrophic bacteria. J Ind Microbiol Biotechnol **33:**88-93.

136. **Cogdell, R. J., T. D. Howard, R. Bittl, E. Schlodder, I. Geisenheimer, and W. Lubitz.** 2000. How carotenoids protect bacterial photosynthesis. Philos Trans R Soc Lond B Biol Sci **355:**1345-9.

137. **Plonka, P. M., and M. Grabacka.** 2006. Melanin synthesis in microorganisms--biotechnological and medical aspects. Acta Biochim Pol **53:**429-43.

138. **Alvarez-Martinez, C. E., R. L. Baldini, and S. L. Gomes.** 2006. A *Caulobacter crescentus* extracytoplasmic function sigma factor mediating the response to oxidative stress in stationary phase. J Bacteriol **188:**1835-46.

139. **Mostertz, J., C. Scharf, M. Hecker, and G. Homuth.** 2004. Transcriptome and proteome analysis of *Bacillus subtilis* gene expression in response to superoxide and peroxide stress. Microbiology **150:**497-512.

140. **Moradas-Ferreira, P., V. Costa, P. Piper, and W. Mager.** 1996. The molecular defences against reactive oxygen species in yeast. Mol Microbiol **19:**651-8.

141. **Chudakov, D. M., S. Lukyanov, and K. A. Lukyanov.** 2005. Fluorescent proteins as a toolkit for in vivo imaging. Trends Biotechnol **23:**605-13.

142. **Habuchi, S., R. Ando, P. Dedecker, W. Verheijen, H. Mizuno, A. Miyawaki, and J. Hofkens.** 2005. Reversible single-molecule photoswitching in the GFP-like fluorescent protein Dronpa. Proc Natl Acad Sci U S A **102:**9511-6.

143. **Gurskaya, N. G., V. V. Verkhusha, A. S. Shcheglov, D. B. Staroverov, T. V. Chepurnykh, A. F. Fradkov, S. Lukyanov, and K. A. Lukyanov.** 2006. Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. Nat Biotechnol **24:**461-5.

144. **Merzlyak, E. M., J. Goedhart, D. Shcherbo, M. E. Bulina, A. S. Shcheglov, A. F. Fradkov, A. Gaintzeva, K. A. Lukyanov, S. Lukyanov, T. W. Gadella, and D. M. Chudakov.** 2007. Bright monomeric red fluorescent protein with an extended fluorescence lifetime. Nat Methods **4:**555-7.

145. **Ai, H. W., J. N. Henderson, S. J. Remington, and R. E. Campbell.** 2006. Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging. Biochem J **400:**531-40.

146. **Shaner, N. C., R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer, and R. Y. Tsien.** 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. Nat Biotechnol **22:**1567-72.

147. **Tsien, R. Y.** 1998. The green fluorescent protein. Annu Rev Biochem **67:**509-44.

148. **Losi, A., E. Ghiraldelli, S. Jansen, and W. Gärtner.** 2005. Mutational effects on protein structural changes and interdomain interactions in the blue-light sensing LOV protein YtvA. Photochem Photobiol **81:**1145-52.

149. **Johansson, L. B., A. Davidsson, G. Lindblom, and K. R. Naqvi.** 1979. Electronic transitions in the isoalloxazine ring and

orientation of flavins in model membranes studied by polarized light spectroscopy. Biochemistry **18:**4249-53.

150. **Klipp, W., B. Masepohl, and A. Puhler.** 1988. Identification and mapping of nitrogen fixation genes of *Rhodobacter capsulatus*: duplication of a nifA-nifB region. J Bacteriol **170:**693-9.

151. **Shimomura, O., F. H. Johnson, and Y. Saiga.** 1962. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. J Cell Comp Physiol **59:**223-39.

152. **Sidhu, S. S., and A. A. Kossiakoff.** 2007. Exploring and designing protein function with restricted diversity. Curr Opin Chem Biol **11:**347-54.

153. **Ward, W. H., D. Timms, and A. R. Fersht.** 1990. Protein engineering and the study of structure--function relationships in receptors. Trends Pharmacol Sci **11:**280-4.

154. **Wells, J. A.** 1991. Systematic mutational analyses of protein-protein interfaces. Methods Enzymol **202:**390-411.

155. **Wells, J. A.** 1996. Binding in the growth hormone receptor complex. Proc Natl Acad Sci U S A **93:**1-6.

156. **Weng, L. X., L. H. Wang, J. L. Xu, J. E. Wu, Q. Li, and L. H. Zhang.** 2005. Molecular and conformational basis of a specific and high-affinity interaction between AlbA and albicidin phytotoxin. Appl Environ Microbiol **71:**1445-52.

157. **Clackson, T., and J. A. Wells.** 1995. A hot spot of binding energy in a hormone-receptor interface. Science **267:**383-6.

158. **Cunningham, B. C., D. J. Henner, and J. A. Wells.** 1990. Engineering human prolactin to bind to the human growth hormone receptor. Science **247:**1461-5.

159. **Cunningham, B. C., and J. A. Wells.** 1989. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science **244:**1081-5.

160. **Buttani, V., A. Losi, E. Polverini, and W. Gärtner.** 2006. Blue news: NTP binding properties of the blue-light sensitive YtvA protein from *Bacillus subtilis*. FEBS Lett **580:**3818-22.

161. **Kjeldgaard, M., J. Nyborg, and B. F. Clark.** 1996. The GTP binding motif: variations on a theme. Faseb J **10:**1347-68.

162. **Leipe, D. D., Y. I. Wolf, E. V. Koonin, and L. Aravind.** 2002. Classification and evolution of P-loop GTPases and related ATPases. J Mol Biol **317:**41-72.

163. **Najafi, S. M., D. A. Harris, and M. D. Yudkin.** 1996. The SpoIIAA protein of *Bacillus subtilis* has GTP-binding properties. J Bacteriol **178:**6632-4.

164. **Zhang, S., and W. G. Haldenwang.** 2005. Contributions of ATP, GTP, and redox state to nutritional stress activation of the *Bacillus subtilis* sigmaB transcription factor. J Bacteriol **187:**7554-60.

165. **Piggot, P. J., and D. W. Hilbert.** 2004. Sporulation of *Bacillus subtilis*. Curr Opin Microbiol **7:**579-86.

166. **Yudkin, M. D., and J. Clarkson.** 2005. Differential gene expression in genetically identical sister cells: the initiation of sporulation in *Bacillus subtilis*. Mol Microbiol **56:**578-89.

167. **Carniol, K., T. J. Kim, C. W. Price, and R. Losick.** 2004. Insulation of the sigmaF regulatory system in *Bacillus subtilis*. J Bacteriol **186:**4390-4.

168. **Baird, G. S., D. A. Zacharias, and R. Y. Tsien.** 2000. Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. Proc Natl Acad Sci U S A **97:**11984-9.

169. **Remington, S. J.** 2002. Negotiating the speed bumps to fluorescence. Nat Biotechnol **20:**28-9.

170. **Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, D. A. Zacharias, and R. Y. Tsien.** 2002. A monomeric red fluorescent protein. Proc Natl Acad Sci U S A **99:**7877-82.

7 Appendix

Appendix A: Chemical structures of flavins.

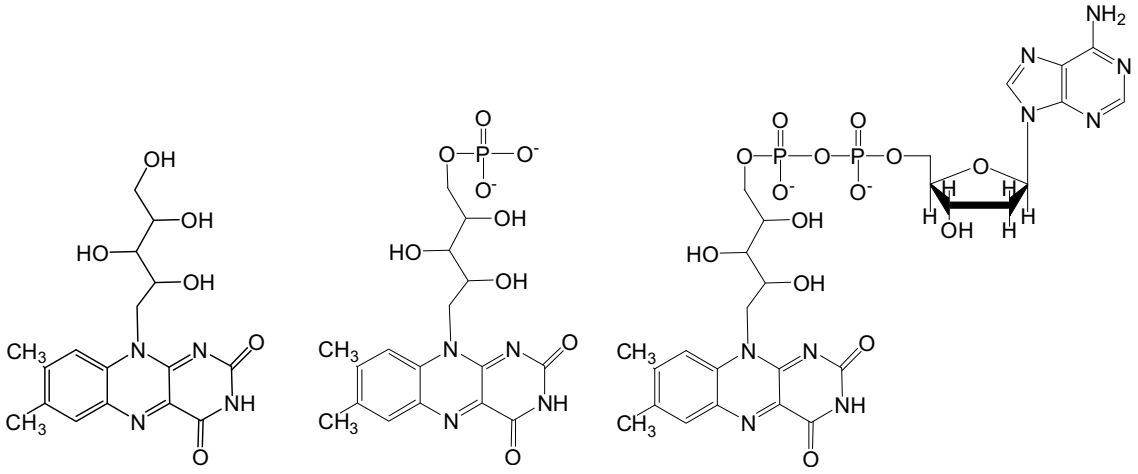


Figure A1: Chemical structures different flavins: riboflavin (RF), flavin mononucleotide (FMN) and Flavin adenine dinucleotide (FAD).

Appendix B: Exemplary chemical structure of a pyoverdine-siderophore.

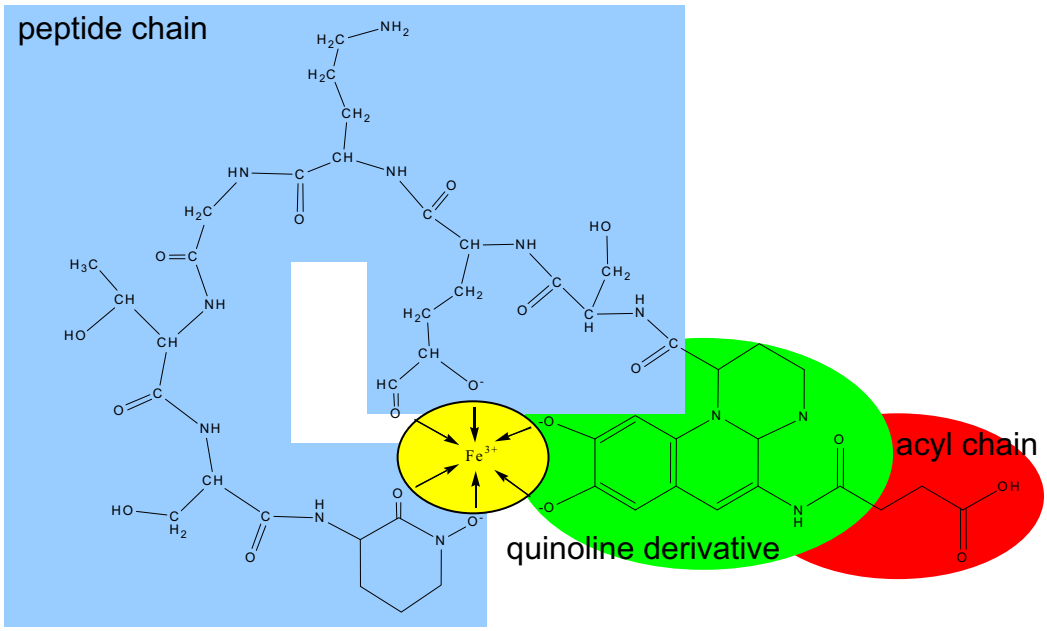


Figure B1: Illustration of the chemical structure of a typical pyoverdine(pvd). Exemplary the pyoverdine of *Pseudomonas aeruginosa* ATCC 27853 is depicted. The ferrous-iron is chelated with high affinity by catecholate, hydroxamate and sometimes β-hydroxy acid groups in the pyoverdine-siderophore. Highlighted are the structural building blocks that constitute the ferrous iron chelation sphere (in yellow) of pyoverdines. Those three buildings blocks are: i) a conserved dihydroxyquinoline chromphore whose conjugated π-ring system renders the pyoverdines intrinsically fluorescent (in green). ii) an acyl side chain (either dicarboxylic acid or amide) that is coupled to the amino group of the quinoline-chromophore (in red), and iii) a variable peptide chain linked by an amide group bound to the C1 (or rarely C3) carboxyl group of the chromophore (in blue).

Supplementary Material: **TABLE 1: Overview of the putative Light, Oxygen, Voltage (LOV) domain containing proteins**

Bacteria: α-Proteobacteria

Species	16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Rhodobacter sphaeroides 2.4.1	Alpha1	Q3J4A0	Alpha1	LOV
Dinoroseobacter shibae DFL 12 (1)	Alpha2	ZP_01584778, A0VR44	Alpha2_1	LOV+Kinase
(2)		ZP_01586748	Alpha2_2	LOV+PAS+PAS+Kinase+RR
Fulvimarina pelagi HTCC2506	Alpha3	ZP_01438521, Q0G496	Alpha3	LOV+KInase
Aurantimonas sp. SI85-9A1 (1)	Alpha4	Q1YFS4	Alpha4_1	LOV+Kinase
(2)		Q1YEU2	Alpha4_2	LOV+Kinase
Oceanicola granulosus HTCC2516	Alpha5	Q2CIF5	Alpha5	LOV+Kinase
Roseobacter denitrificans OCh 114	Alpha6	Q167W8	Alpha6	LOV+Kinase
Parvularcula bermudensis HTCC2503 (1)	Alpha7	ZP_01016632 , A3VPU8	Alpha7_1	LOV+Kinase
(2)		ZP_01018013, A3VTS2	Alpha7_2	LOV+Kinase
Erythrobacter litoralis HTCC2594 (1)	Alpha8	Q2NB98	Alpha8_1	LOV+HTH
(2)		Q2NB77	Alpha8_2	LOV+Kinase
(3)		Q2N9L9	Alpha8_3	LOV+Kinase
(4)		Q2NCA3	Alpha8_4	LOV+Kinase
Novosphingobium aromaticivorans DSM12444 (1)	Alpha9	Q2G5U0	Alpha9_1	LOV+Kinase
(2)		Q2G8Z7	Alpha9_2	LOV+HTH
Brucella melitensis 16M	Alpha10	Q8YC53	Alpha10	LOV+PAS+Kinase
Brucella abortus biovar 1 str. 9-941	Alpha11	Q577Y7	Alpha11	LOV+PAS+Kinase
Brucella suis	Alpha12	Q8FW73	Alpha12	LOV+PAS+Kinase
Caulobacter crescentus CB15	Alpha13	Q9ABE3	Alpha13	LOV+Kinase
Caulobacter sp. K31 (1)	Alpha14	Q0LTE1	Alpha14_1	LOV+Kinase
(2)		Q0M3Z0	Alpha14_2	LOV+Kinase
Magnetospirillum magnetotacticum MS-1 (1)	Alpha15	ZP_00051334	Alpha15_1	LOV+Kinase
(2)		ZP_00052303	Alpha15_2	LOV+Kinase
Sphingopyxis alaskensis RB2256	Alpha16	Q1GUF5	Alpha16	LOV+HTH
Sphingomonas sp. SKA58 (1)	Alpha17	Q1N7J1	Alpha17_1	LOV+HTH
(2)		Q1NI33	Alpah17_2	LOV+LOV+Kinase+RR
Bradyrhizobium sp. BTAi1	Alpha18	Q35E64	Alpha18	LOV+Kinase+RR
Rhizobium leguminosarum bv viciae 3841 plasmid pRL11	Alpha19	Q1M667	Alpha19	LOV+Kinase

Bacteria: β-Proteobacteria

Species	16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Burkholderia phymatum STM815	Beta1	A0FWN5	Beta_1	LOV+Kinase+RR
Acidovorax avenae sub. citrulli AAC00-1	Beta2	Q0XDC0	Beta_2	LOV+Kinase+RR
Nitrosospira multiformis ATCC 25196	Beta3	Q2Y837	Beta_3	LOV
Ralstonia solanacearum UW551 (1)	Beta4	CAD17405	Beta4_1	HAMP+PAS+PAS+LOV+GGDEF+EAL
Ralstonia solanacearum GMI1000 (2)		Q8XT61	Beta4_2	HAMP+PAS+PAS+LOV+GGDEF+EAL
Rubrivivax gelatinosus PM1 (Typestr.16S)	Beta5	ZP_00244844	Beta5	GAF+PAS+PAS+PAS+LOV+PAS+PAS+Kinase+RR
Polaromonas naphthalenivorans	Beta6	ZP_01020075	Beta6	LOV+Kinase
Burkholderia xenovorans LB400	Beta7	Q145M2	Beta7	PAS+PAS+LOV+GGDEF+EAL

Bacteria: γ-Proteobacteria

Species	16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Xanthomonas axonopodis pv. citri str.306	Gamma1	Q8PJH6	Gamma1	LOV+Kinase+RR
Xanthomonas campestris pv campestris ATCC 33913	Gamma2	Q8P827	Gamma2	LOV+Kinase+RR
Xanthomonas oryzae pv. oryzae MAFF 311018	Gamma3	Q2P134	Gamma3	LOV+Kinase+RR
Xanthomonas campestris pv. Vesicatoriastr.85-10	Gamma4	Q3BRX8	Gamma4	LOV+Kinase+RR
Pseudomonas syringae pv. syringae B728a	Gamma5	Q4ZSY3	Gamma5	LOV+Kinase+RR
Pseudomonas syringae pv. tomato DC3000	Gamma6	Q881J7	Gamma6	LOV+Kinase+RR
Pseudomonas syringae pv. Phaseolicola1448A	Gamma7	Q48IV1	Gamma7	LOV+Kinase+RR
Nitrosococcus oceani ATCC 19707	Gamma8	Q3J6W8	Gamma8	LOV
Thiomicrospira denitrificans ATCC 33889	Gamma9	Q30NS0	Gamma9	HTH+LOV
Alteromonas macleodii 'Deep ecotype'	Gamma10	ZP_01109417	Gamma10	LOV+Kinase+RR
Pseudomonas fluorescens PF0-1	Gamma11	Q3KHW7	Gamma11	LOV
Pseudomonas fluorescens PF-5	Gamma12	Q4BW45	Gamma12	LOV
Pseudomonas putida KT2440 (1)	Gamma13	Q88E39	Gamma13_1	LOV
(2)		Q88JB0	Gamma13_2	LOV
Pseudomonas putida F1 (1)	Gamma14	ZP_00900893	Gamma14_1	LOV
(2)		ZP_00898075	Gamma14_2	LOV
Pseudomonas putida W619 (1)	Gamma15	ZP_01637579	Gamma15_1	LOV
(2)		ZP_01638519	Gamma15_2	LOV
Chromohalobacter salexigens DSM 3043	Gamma16	Q1QU87	Gamm16	LOV

Bacteria: Chloroflexi

Species	16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Chloroflexus aurantiacus	Chloro1	Q3DVF5	Chloro1	LOV+PAS+PAS+Kinase
Chloroflexus aggregans DSM 9485	Chloro2	A0H5C8	Chloro2	LOV+PAS+PAS+Kinase+RR
Herpetosiphon aurantiacus ATCC 23779	Chloro3	ZP_01426530	Chloro3	HAMP+GAF+LOV+PAS+PAS+PAS+PAS+GAF+Kinase

Bacteria: Actinobacteria

Species	16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Kineococcus radiotolerans (1) (2)	Actino1	Q40X75 Q40XK1	Actino1_1 Actino1_2	GGDEF+LOV+EAL LOV+GAF+SpoIIE
Rubrobacter xylanophilus	Actino2	Q1ARZ9	Actino2	LOV+GAF+SpoIIE

Bacteria: Firmicutes

Species	16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Listeria monocytogenes EGD-e	Firmi1	P58724	Firmi1	LOV+STAS
Listeria monocytogenes F2365	Firmi2	Q722B8	Firmi2	LOV+STAS
Listeria innocua CLIP 11262	Firmi3	Q92DM1	Firmi3	LOV+STAS
B.subiilis 168	Firmi4	O34627	Firmi4	LOV+STAS
Oceanobacillus iheyensis HTE831	Firmi5	Q8ESN8	Firmi5	LOV+STAS
Listeria *monocytogenes* str. 1/2a F6854	Firmi6	ZP_00235050.1	Firmi6	LOV+STAS
Bacillus amyloliquefaciens FZB42	Firmi7	YP_001422316.1	Firmi7	LOV+STAS
Listeria welshimeri serovar 6b str. SLCC5334	Firmi8	YP_848958.1	Firmi8	LOV+STAS

Bacteria: Cyanobacteria

Species		16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Thermosynechococcus elongatus BP-1		Cyano1	Q8DJE3	Cyano1	GAF+PAS+LOV+GAF+GAF+Kinase+RR
Anabaena variabilis ATCC 29413	(1)	Cyano2	Q3M6B3	Cyano2_1	LOV+PAS+PAS+GGDEF+EAL
	(2)		Q3MED3	Cyano2_2	PAS+LOV+PAS+PAS+GAF+Kinase+RR+RR+HPT
Synechococcus sp. PCC 6301	(1)	Cyano3	Q5N2F7	Cyano3_1	LOV+GGDEF+EAL
	(2)			Cyano3_2	RR*PAS+PAS+LOV+GGDEF+EAL
Nostoc punctiforme PCC 73102	(1)	Cyano4	ZP_00111211	Cyano4_1	LOV+PAS+PAS+GGDEF+EAL
	(2)		ZP_00105980	Cyano4_2	PAS+LOV+PAS+GAF+Kinase+RR+RR+HPT
Crocosphaera watsonii		Cyano5	Q4BW45	Cyano5	RR+LOV
Lyngbya sp. PCC 8106	(1)	Cyano6	ZP_01623058	Cyano6_1	RR+LOV+GGDEF
	(2)		ZP_01624062	Cyano6_2	PAS+LOV+PAS+GGDEF+EAL
	(3)		ZP_01621849	Cyano6_3	RR+RR+LOV+PAS+GAF+Kinase+RR
	(4)		ZP_01618684	Cyano6_4	ÜAS+LOV+PAS+PAS+PAS+GAF+GGDEF
Cyanothece sp. CCY0110		Cyano8	ZP_01728401.1	Cyano8	RR+LOV+DUF
Anabaena sp PCC7120*(Nostoc)*	(1)	Cyano9	all2875 (Cyanobase ID)	Cyano9_1	LOV+PAS+PAS+DUF1+DUF2
	(2)		alr3170 (Cyanobase ID)	Cyano9_2	PAS+LOV+PAS+PAS+PAS+GAF+Kinase+RR

Eukaryota: Viridiplantae

Species		16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Physcomitrella patens	(1)	Plant1	BAD32622	Plant1_1L1 / Plant1_1L2 (phot1)	LOV1+LOV2+Kinase
	(2)		BAD32623	Plant1_2L1 / Plant1_2L2 (phot2)	LOV1+LOV2+Kinase
	(3)		BAD32624	Plant1_3L1 / Plant1_3L2 (phot3)	LOV1+LOV2+Kinase
Oryza sativa	(1)	Plant2	XP_471720	Plant2L1 / Plant2L2 (phot)	LOV1+LOV2+Kinase
	(2)		Q5Z8K3	Plant2_3 (ZTL)	LOV+Fbox+$Kelch_5$
	(3)		Q67UX0	Plant2_4 (ZTL)	LOV+Fbox+$Kelch_5$
Chlamydomonas reinhardtii		Plant3	CAC94940	Plant3L1 / Plant3L2 (phot)	LOV1+LOV2+Kinase
Arabidopsis thaliana	(1)	Plant4	NP_849983	Plant4_1	LOV+Fbox+$Kelch_5$
	(2)		Q94BT6	Plant4_2	LOV+Fbox+$Kelch_5$
	(3)		AAK27435	Plant4_3	LOV+Fbox+$Kelch_5$
	(4)		Q9C9W9	Plant4_4	LOV+Fbox+$Kelch_5$
	(5)		NM_114447.3	Plant4P1L1 / Plant4P1L2 (phot1)	LOV1+LOV2+Kinase
	(6)		NM_180881.2	Plant4P2L1 / Plant4P2L2 (phot2)	LOV1+LOV2+Kinase
Mesembryanthemum crystallinum	(1)	Plant5	Q6UEI5	Plant5_1	LOV+Fbox+$Kelch_5$
	(2)		Q6UEI4	Plant5_2	LOV+Fbox+$Kelch_5$
	(3)		gi|457692(LOV2-Kinase)	Plant5_3	LOV1+LOV2+Kinase

Eukayrota: Fungi

Species		16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Botrytis cinerea B05.10	(1)	Fungi1	BC1G_13505.1	Fungi1_1	LOV+PAS+PAS+ZnF
	(2)		BC1G_08584.1	Fungi1_2	RGS+LOV
	(3)		BC1G_15404.1	Fungi1_3	LOV
Sclerotinia sclerotiorum	(1)	Fungi2	SS1G_02240	Fungi2_1	LOV
	(2)		SS1G_11953	Fungi2_2	LOV+PAS+PAS+ZnF
	(3)		SS1G_12563	Fungi2_3	LOV
Magnaporthe grisea 70-15	(1)	Fungi3	MGG_03538.5	Fungi3_1	LOV+PAS+PAS+ZnF
	(2)		MGG_01041.5	Fungi3_2	LOV
	(3)		MGG_08735.5	Fungi3_3	RGS+LOV
Aspergillus nidulans	(1)	Fungi4	AN3436.1	Fungi4_1	LOV+PAS+PAS+ZnF
	(2)		AN3435	Fungi4_2	LOV
Rhizopus oryzae	(1)	Fungi5	RO3G_14273.1	Fungi5	LOV+PAS+ZnF
Neurospora crassa		Fungi6	NCU02356	Fungi6	LOV+PAS+PAS+ZnF
Cryptococcus neoformans		Fungi7	Q6DMM2	Fungi7	LOV+PAS

Archaea: Euryarcheota

Species		16s Tree-ID	Accession Number of LOV protein	LOV protein Tree-ID	LOV protein domain content
Haloarcula marismortui ATCC 43049	(1)	Arch1	Q5V5P7	Arch1_1	RR+PAS+LOV+PAS+Kinase
	(2)		Q5V3S3	Arch1_2	RR+PAS+LOV+GAF+Kinase
	(3)		Q5UWI7	Arch1_3	PAS+GAF+PAS_4+GAF+PAS_4+LOV+PAS_3+Linase
Natronomonas pharaonis DSM 2160	(1)	Arch2	Q3ITW5	Arch2_1	RR+LOV+GAF+HTH
plasmid PL131	(2)		Q3IM51	Arch2_2	PAS+LOV+PAS+Kinase
Halorubrum lacusprofundi ATCC 49239	(1)	Arch3	ZP_02014689.1	Arch3_1	PAS+LOV+Kinase
	(2)		EDN50126.1	Arch3_2	PAS+LOV+Kinase

8 Acknowledgements

First of all I would like to thank Prof. Dr. Karl-Erich Jaeger[A](at the Institut für Molekulare Enzymtechnologie in Jülich) and Prof. Dr. Wolfgang Gärtner[B] (at the Max-Planck-Institute for Bioinorganic Chemisty, in Mülheim an der Ruhr) for providing me with this challenging and up-to-date topic and moreover, for taking over the tedious task of writing the reviews. Furthermore, I would particularly like to express my gratitude to them for taking such great interest in my work, the numerous inspiring debates as well as the continuous help and guidance with the publication of the results of this thesis.

Very special thanks go to my supervisor on-site, PD Dr. Thorsten Eggert[C] for the overall opportunity to work in his group and of course for the many helpful and encouraging discussions. Moreover, I am especially grateful that he provided me with the freedom to independently pursue my ideas while at the same time taking care of the necessary critical review, and therefore not letting me stray to deep into the nebulous territories of speculation. Furthermore, his guidance and support during the sometimes cumbersome task of writing my papers was highly valued and appreciated.

Herewith, I would like to thank the Deutsche Forschungsgemeinschaft (DFG) for the financial support of my work under the framework of the DFG-Forschergruppe 526 “Blue-Light Photoreceptors”.

Every journey is easier when you travel together. Therefore, I would like to thank my following collaborators. Without them and their continuous dedication this work would not have been possible.

First of all my very special thanks go to Dr. Aba Losi[D] (Department of Physics, University of Parma, Italy) for the ongoing fruitful cooperation, many helpful discussions, her advice on certain aspects of biophysics and last but not least for her critical review of my thesis and papers. She certainly was and is a role model for me, embodying how commitment to science should look like.

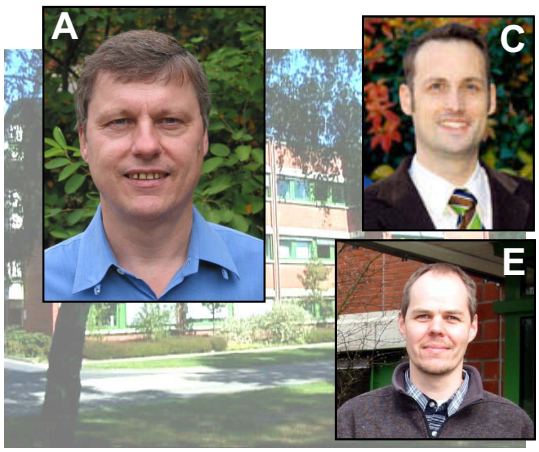
Many thanks go to Dr. Thomas Drepper[E] (AG Bacterial Photobiotechnology, Institut für Molekulare Enzymtechnologie, in Jülich), for the many inspiring debates regarding bacterial physiology, his help with the generation of knockout-mutants and sharing of plasmid-constructs, his critical reading of parts of my thesis and lastly also for the space he provided me with in his laboratory, a place with a friendly and blithe working atmosphere under the “*Rhodobacter*-Sun”.

Very special thanks also to Prof. Dr. Arndt von Haeseler[F] (Center for Integrative Bioinformatics, Max F. Perutz Laboratories, Vienna, Austria), for working with us on the LOV-phylogeny project, whose efforts and timely assistance in finishing the “LOV is all around” manuscript allowed me to submit my thesis in time. In the same context, I would like to thank his PhD student Minh Bui Quang[G] for introducing me to the possibilities and limitations of up-to-date phylogentic tree reconstruction methods

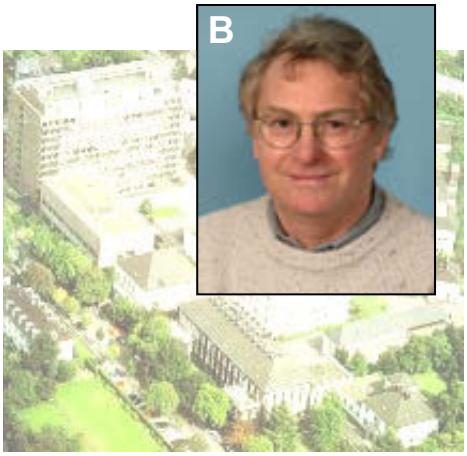
Thank you very much also to Rita Reichel and the team of inlingua Chemnitz for the punctual assistance in putting the finishing touches to my thesis.

Furthermore, loads of thanks to all the former and present members of the AG Eggert and AG Drepper who shared with me lab, office, coffee, a laugh and a relaxed working atmosphere. Just to mention a few: Solmaz Arvani, Vera Svenson, Benjamin Franken, Franco Circolone, Michael Puls, Jennifer Andexer, Jan-Karl Guterl, Eliane Bogo … .In particular I would like to thank Astrid Wirtz for her help with the generation of various mutants and her excellent assistance in setting-up and performing HPLC analysis on all of my LOV proteins.

Last but not least I would like to thank my parents[H] for their faith in me as well as for their support and assistance during all my studies. Finally, very special thanks to Karen[I], for giving me love, friendship, help and strength. She has always been close to me, both through successes and difficult times, especially during the last demanding months of lab-work and writing, when I was sometimes engrossed in thoughts or lingered around, literally speaking, in other spheres. Sorry, my love. It is done!



Institut für Molekulare Enzymtechnologie, Jülich

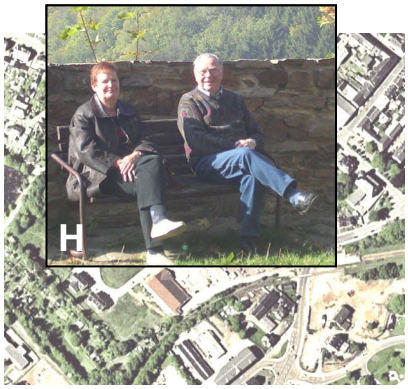


Max-Planck Institute, Mülheim a.d. Ruhr

Universita degli studi di Parma, Departimento di Fisica, Parma, Italia

Max F.Perutz Laboratories, Center for Integrative Bioinformatics, Wien, Österreich

Aachen, Nordrhein-Westfalen



Marienberg, Sachsen

A: Prof. Dr. Karl-Erich Jaeger, B: Prof. Dr. Wolfgang Gärtner, C: PD Dr. Thorsten Eggert, D: Dr. Aba Losi, E: Dr. Thomas Drepper, F: Prof. Dr. Arndt von Haeseler, G: Minh Bui, Quang, H: Karl und Marianne, I,J: Karen und Luna

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 09.11.2007