

Structure and dynamics of a blue-light activated LOV photoreceptor

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Raj Rani
from Nangal (India)

Jülich, November 2012

from the institute for Molekulare Enzymtechnologie (IMET)
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Karl-Erich Jaeger
Co-supervisor: Prof. Dr. Dieter Willbold

Date of the oral examination: 18.12.12

Acknowledgment

Every piece of creation is originated by zeal of hard work and determination, and to put the efforts into action, many factors influence. Nothing concrete can be achieved without inspiration and perspiration.

Any work is rarely the product of only one person, and this dissertation is absolutely no exception. It gives me immense pleasure to express my deepest gratitude towards all those persons who have contributed actively, their exceptional diligence and contribution to this project are highly appreciated. First and foremost I would like to thank my excellent thesis advisors, Prof. Karl-Erich Jaeger (Institut für Molekulare Enzymtechnologie in Jülich) and Prof. Dieter Willbold (Institute of Complex Systems in Jülich) for giving me the great opportunity to pursue a doctorate in this very interesting field. I am glad I persevered and I am grateful for their guidance in my formative years. Without their vision and support I would not have been able to reach this milestone.

I owe my special gratitude to Dr. Ulrich Krauss (at the Institut für Molekulare Enzymtechnologie in Jülich) for excellent guidance and invaluable suggestions over the years. I appreciate and thank him for all the support as well as for critically reading this thesis.

I want to thank Dr. Rudolf Hartmann (at the Institute of Complex Systems in Jülich) for introducing me to NMR spectroscopy, measuring NMR experiments and for building up the fiber optic illumination system used during my project work.

I would also like to thank Dr. Justin Lecher and Dr. Melanie Schwarten for their help with NMR assignments and structure calculations. Special thanks to Justin Lecher for his help with calculation of SSP scores, peptide structure calculations and database submissions. I appreciate all the support they have given me during endless discussions.

I would like to thank Esther Knieps-Grünhagen and Maya Piqueray for their continuous support i.e. by providing laboratory materials and assistance. Thanks also to Astrid Wirtz for HPLC analyses.

Now I thank my colleagues Kathrin Scholz, Katrin Jentzsch, Marco Kaschner, Benita Kopka and Martin Diener for creating the friendly, cooperative environment in our lab. I also thank my friends Sonia Biswas, Hoa Quynh Do and Franco Circolone for unforgettable moments and their everlasting support.

I would like to thank the whole IMET and ICS-6 teams for providing such a close and friendly environment to work.

I would like to thank NRW- Research School BioStruct for their financial support for this project work. Special thanks to Christian Dumpitak, Cordula Kruse and Brigitte Haumann, they made official procedures very easy.

Last but not least, I wish to express my gratitude to my most beloved and respected parents who occupy the top seat in my world, for their benign blessing and constant moral.

I want to dedicate this thesis to my parents, my husband Gaurav and my daughter Devshi. I owe my husband for his unselfish support and forever love which enabled me to focus on my target.

Finally I thank the almighty for giving me an opportunity to thank all these people.

Publication and Poster presentations

Raj Rani, Katrin Jentsch, Justin Lecher, Rudolf Hartmann, Dieter Willbold, Karl-Erich Jaeger, Ulrich Krauss (2013).

Conservation of dark recovery kinetic parameters and structural features in the *Pseudomonadaceae* “short” LOV protein family – implications for the design of LOV-based optogenetic tools. *Biochemistry*. 52(26):4460-4473

Poster presentations

Raj Rani, Karl-Erich Jaeger, Dieter Willbold and Ulrich Krauss

“Structural studies on the twin LOV-proteins of *Pseudomonas putida*” in DFG-FOR 526 “Blue-light photoreceptors” meeting, March 2010, Frauenchiemsee, Germany

Raj Rani, Rudolf Hartmann, Katrin Jentsch, Karl-Erich Jaeger, Dieter Willbold and Ulrich Krauss

“Solution NMR spectroscopy to study light dependent conformational changes in one LOV photoreceptor protein of *Pseudomonas putida*”, Biostruct master class, September 2010, Duesseldorf, Germany

Raj Rani, Rudolf Hartmann, Karl-Erich Jaeger, Dieter Willbold and Ulrich Krauss

“Solution NMR spectroscopy to study chemical shift perturbation upon blue light illumination in the PpSB1-LOV photoreceptor protein of *Pseudomonas putida*”, Biostruct master, September 2011, Duesseldorf, Germany

Raj Rani, Rudolf Hartmann, Katrin Jentsch, Justin Lecher, Karl-Erich Jaeger, Dieter Willbold and Ulrich Krauss

“Conservation of coiled-coil structural motifs in the “short” LOV protein family” Gordon research conference: Photosensory Receptors and Signal Transduction, January 2012, Galveston, Texas, USA

Abstract

Light, Oxygen, Voltage (LOV) domains are part of ubiquitously distributed blue-light activated photoreceptors that carry non-covalently bound flavin mononucleotide (FMN) as the light sensitive chromophore to capture photons of the incident radiation. Upon blue-light illumination, LOV domains undergo a spectroscopically but not structurally well characterized photocycle. Based on spectroscopic and X-ray crystallographic data for several plant and bacterial LOV domains, it was suggested that blue-light illumination results in the formation of a covalent bond between a totally conserved cysteine residue and the C4a carbon atom of the flavin isoalloxazine ring. So far little evidence for large-scale light dependent conformational changes associated with photon capture in LOV domains could be provided by X-ray crystallography probably due to the restraint imposed by the presence of the crystal-lattice. On the other hand solution methods such as Nuclear Magnetic Resonance (NMR) and Circular Dichroism (CD) spectroscopy used to study the oat phototropin1-LOV2 domain and other plant LOV domains revealed light-dependent conformational changes in the sensor domain which in turn are probably relayed to fused effector domains via the dissociation or unfolding of a C-terminally located α -helical linker-segment (termed $J\alpha$ -helix). This eventually results in activation of the effector. So far, for bacterial LOV systems, such as the short bacterial LOV proteins *Pseudomonas putida* studied here, no light-dependent loss of helical content or other widespread secondary structural changes have been observed.

Here we exemplarily use the LOV-proteins PpSB1-LOV and PpSB2-LOV of *Pseudomonas putida* as a paradigm for an extended LOV-construct. PpSB1-LOV exhibits a plant phototropin LOV-like photochemistry. The protein consists of a conserved LOV core flanked by short N- and C-terminal extensions but lacks a fused effector domain. CD spectroscopy indicates that both LOV-core extensions are helical in similarity to the oat phototropin1-LOV2 construct. However, as for other bacterial LOV proteins, a clear loss of helical content (or any other changes in secondary structure) in response to blue-light illumination has so far not been observed. In order to elucidate the role (and the structure) of the respective extensions, we generated truncated constructs of PpSB1-LOV and PpSB2-LOV. To our surprise, the truncation of both the N- and

C-terminal extensions results in a complete loss of solubility of the respective overexpressed protein, indicating that both regions contribute to the proper folding and structural integrity. To investigate the folding of the isolated C-terminal extension the respective polypeptide was synthesized and studied by CD-spectroscopy and $^1\text{H-NMR}$ techniques. Both methods suggest that PpSB1-J α and PpSB2-J α peptides adopt an α -helical conformation in solution.

In order to understand the light-dependent conformational changes in the full-length photoreceptor protein, we used NMR spectroscopy to structurally characterize the dark- and light-state of the protein in solution. After optimization of expression and NMR buffer conditions, we obtained a complete backbone resonance assignment of both the dark- and light-state of PpSB1-LOV. Upon illumination we observe pronounced chemical shift and peak intensity changes for most backbone amide resonances. The data indicate severe light-dependent conformational changes which so far have not been observed for the PpSB1-LOV protein by using other solution techniques. The near complete backbone assignment enabled us to use various types of NMR-based experiments such as H/D exchange and relaxation experiments to characterize the protein dynamics in both states. Those experiments provide valuable insight into the principles of LOV photo-activation, intra-molecular signal relay and signal transduction for this class of bacterial LOV photoreceptors and, at the same time, could guide the structure-based design of LOV-based genetically encoded photoswitches.

Zusammenfassung

Light, *Oxygen*, *Voltage* (LOV) Domänen fungieren als Sensormodule in einer Vielzahl von Blaulicht Photorezeptoren in allen drei Reichen des Lebens (*Eukarya*, *Prokarya* und *Archaea*). Sie binden Flavinmononukleotid (FMN) als lichtsensitiven (absorbierenden) Kofaktor oder Chromophor. Alle bisher beschriebenen LOV Photorezeptoren durchlaufen bei Beleuchtung mit blauem Licht einen spektroskopisch gut, jedoch strukturell unvollständig, verstandenen Photozyklus. Basierend auf spektroskopischen und röntgenkristallographischen Daten konnte gezeigt werden, dass eine Beleuchtung mit Blaulicht zur Ausbildung einer kovalenten Bindung zwischen dem C4a Atom des FMN Chromophores und einem konserviertem Cysteinrest in der LOV Domäne führt (Lichtzustand). Im Dunkeln wird diese Bindung gebrochen, wobei das LOV Protein in seinen Dunkelzustand zurückkehrt. In den entsprechenden Kristallstrukturen konnten jedoch keine größeren, lichtabhängigen Veränderungen nachgewiesen werden. Eine Ursache hierfür könnte in der eingeschränkten Beweglichkeit des Proteinmoleküls zu suchen sein, die durch die Kristallpackung hervorgerufen wird. Im Gegensatz dazu konnte mittels spektroskopischer Verfahren, wie Kernspin-Resonanz- (*engl. nuclear magnetic resonance, NMR*) und Zirkulardichroismus- (*engl. circular dichroism, CD*) Spektroskopie gezeigt werden, dass LOV Domänen pflanzlicher LOV Photorezeptoren (Phototropine) bei Beleuchtung in Lösung dezidierte Strukturänderungen durchlaufen. Insbesondere wurde gezeigt, dass die Signalweiterleitung innerhalb des entsprechenden Photorezeptorproteins wahrscheinlich über die Dissoziation und/oder Entfaltung einer α -helix ($J\alpha$ -helix) am Carboxy-Terminus der LOV Domäne erfolgt. Im Vollängen-Photorezeptor führt dies zur Aktivierung einer fusionierten Serin/Threonin Kinase, welche als sogenannter Effektor oder Signaloutput-Domäne fungiert. Für bakterielle LOV Photorezeptoren konnten bisher keine vergleichbaren Sekundärstrukturänderungen, d.h. insbesondere keine lichtabhängige Entfaltung der $J\alpha$ -helix, beobachtet werden.

In dieser Arbeit wurden exemplarisch zwei bakterielle LOV Proteine, PpSB1-LOV und PpSB2-LOV, aus dem γ -Proteobakterium *Pseudomonas putida* KT2440 untersucht. Beide Proteine durchlaufen bei Beleuchtung mit blauem Licht einen klassischen LOV Photozyklus.

Sie bestehen aus einer konservierten LOV Domäne, welche N- und C-terminal von kurzen Sequenzverlängerungen flankiert wird, besitzen jedoch im Gegensatz zu den meisten anderen LOV Photorezeptoren keine definierte Effektor-domäne. CD-spektroskopische Untersuchungen der beiden Vollängenproteine legen nahe, dass die entsprechenden Verlängerungen, analog zur $J\alpha$ -helix in Phototropin LOV-Domänen, α -helikale Sekundärstrukturen ausbilden. Um die Funktion und Struktur der beiden Verlängerungen zu untersuchen und besser zu verstehen, wurden in dieser Arbeit eine Reihe verkürzter Varianten von PpSB1-LOV und PpSB2-LOV erzeugt. Überraschender Weise führt das Entfernen sowohl der N- als auch der C-terminalen Verlängerung dazu, dass die entsprechenden Varianten in *E. coli* ausschließlich in unlöslicher Form produziert werden. Diese Beobachtung legt nahe, dass die entsprechenden Strukturelemente zur korrekten Faltung und somit zur strukturellen Integrität beider Proteine beitragen. Um die Faltung/Struktur der putativen C-terminalen $J\alpha$ -Helix-Segmente beider Proteine zu validieren, wurden die entsprechenden Polypeptide synthetisiert und mittels CD- und NMR-Spektroskopie untersucht. Beide Methoden zeigen unabhängig voneinander, dass die isolierten $J\alpha$ -Strukturelemente beider Proteine unterschiedlich lange und unterschiedlich stabile, α -Helices in Lösung ausbilden.

Um eine etwaige lichtabhängige Konformationsänderung des PpSB1-LOV Photorezeptors weiter zu untersuchen, wurden in dieser Arbeit NMR-spektroskopische Methoden verwendet. Hierzu wurde das entsprechende Protein sowohl im Dunkelzustand als auch bei Beleuchtung mit blauem Licht in Lösung analysiert. Nach erfolgreicher Optimierung der Expressions-, Reinigungs- und Messbedingungen konnten 99 % der Amidresonanzen des Proteinrückgrates, sowohl im Dunkelzustand als auch bei kontinuierlicher Blaulichtbeleuchtung, in den entsprechenden (^1H - ^{15}N) heteronuklearen Einquantenkohärenz- (für *engl. heteronuclear single quantum coherence, HSQC*) Spektren, zugeordnet werden. Der direkte Vergleich der entsprechenden Spektren des Dunkel- und Lichtzustandes deutet auf ausgeprägte Strukturänderungen des Proteins hin, welche bisher mittels anderer biophysikalischer Methoden, wie CD-Spektroskopie, nicht beobachtet werden konnten. Außerdem erlaubte die vollständige Zuordnung der Amidresonanzen des Proteinrückgrates die Durchführung weiterer NMR-basierter Versuche wie H/D Austausch- und Relaxationsexperimente, welche es erlauben, die Dynamik des Proteins in beiden Zuständen zu untersuchen. Die in dieser Arbeit mittels NMR-Spektroskopie erhaltenen

Erkenntnisse bieten einerseits einen ersten Einblick in die Mechanismen der Lichtaktivierung und Signalweiterleitung von PpSB1-LOV und verwandter Proteine und könnten in der Zukunft das Design artifizierlicher LOV-basierter optogenetischer Schalter zur Kontrolle biologischer Prozesse ermöglichen.

Abbreviations

aa-	Amino acid
AI-FM-	Autoinduction full media
ANN-	Artificial neural network
APS-	Ammonium persulfate
ARIA-	Ambiguous Restraints for Iterative Assignment
ATP/ADP-	Adenosine 5'-triphosphate/ Adenosine 5'-diphosphate
B ₀ -	Applied magnetic field
BLUF-	Blue-Light sensors Using FAD domain
BMRB-	Biological magnetic resonance data bank
bp-	Base pair
bR-	Bacteriorhodopsin
BSA-	Bovine serum albumin
CD-	Circular dichroism
COSY-	Correlation spectroscopy
Cry-	Cryptochrome
C-terminal-	Carboxy-terminal
dH ₂ O-	Distilled water
DNA-	Deoxyribonucleic acid
dNTP-	Deoxyribonucleotide triphosphate
DTT-	Dithiothreitol
EDTA-	Ethylene diamine tetraacetic acid
EMBL-	European molecular biology laboratory
FAD-	Flavin adenine dinucleotide (Riboflavin 5'-adenine diphosphate)
FID-	Free induction decay
FMN-	Flavin mononucleotide (Riboflavin 5'- monophosphate)
FT-	Fourier transform

GAF-	cGMP-specific phosphodiesterases/adenylcyclase/FhlA
GMP-	Guanosine 5'-monophosphate
GTP-	Guanosine 5'-triphosphate
γ -	Nuclear gyromagnetic ratio
H/D-	Hydrogen/deuterium
HD-MM-	High cell density IPTG induction minimal medium
HK-	Histidine kinase domain
HPLC-	High performance liquid chromatography
HSQC-	Heteronuclear single-quantum coherence
HTH-	Helix-turn-helix DNA binding domain
IMAC-	Immobilized metal affinity chromatography
IPTG-	Isopropyl- β -D-thiogalactopyranoside
Kd-	Dissociation constant
kDa -	kilo-Dalton
LB-	Luria Bertani
LOV-	Light, oxygen, voltage domain
LOV390-	LOV light-state
LOV447-	LOV dark-state
LOV660-	LOV excited triplet-state
LOV-HK-	LOV histidine kinase
M-	mol/l
MALDI-TOF-	Matrix-assisted laser desorption/ionization-time-of-flight
MD-	Molecular dynamics
mg-	Milligram
mM-	Millimolar
MW-	Molecular weight
MWCO-	Molecular weight cutoff
N-cap-	Amino-terminal cap

NCBI-	National center for biotechnology information
Ni-NTA-	Nickel-nitrilotriacetic acid
nm-	Nanometer
NMR-	Nuclear magnetic resonance
NOE-	Nuclear Overhauser effect
NOESY-	Nuclear Overhauser effect spectroscopy
N-term-	Amino-terminal
OCP-	Orange carotenoid protein
OD-	Optical density
PAGE-	Polyacrylamide gel electrophoresis
PAS-	Per, Arnt, Sim domain
pCA-	<i>para</i> -coumaric acid
PCR-	Polymerase chain reaction
PDA-	Photodiode array
PDB-	Protein data bank
Pfr-	Far-red absorbing form of phytochrome
Phot-	Phototropin
ppm-	Parts per million (10^{-6})
Pr-	Red absorbing form of phytochrome
PYP-	Photoactive yellow protein
R1-	Longitudinal or spin-lattice relaxation rate
R2-	Transversal or spin-spin relaxation rate
RCI-	Random coil index
rmsd-	Root mean square deviation
rpm-	Rotations per minute
RR-	Response regulator
SDS-	Sodium dodecylsulfate
SEC-	Size-exclusion chromatography

STAS-	Sulfate transporter anti-sigma factor antagonist domain
T1-	Longitudinal or spin-lattice relaxation time
T2-	Transversal or spin-spin relaxation time
TEMED-	N,N,N',N'-tetramethylethylenediamine
TFE-	2,2,2-trifluoroethanol
TMS-	Tetramethylsilane
TOCSY-	Total correlation spectroscopy
Tris-	Tris(hydroxymethyl)-amino-methan
TROSY-	Transverse relaxation optimised spectroscopy
UV-	Ultra-violet
v/v-	Volume per volume
WC-1-	White-collar1 protein
w/v-	Weight per volume
ZnF-	Zinc-finger
ZTL-	ZEITLUPE
ϵ -	molar extinction coefficient
ω -	Larmor frequency
δ -	Chemical shift

All amino acids names are abbreviated using their *one-* (e.g. A- for Alanine) or *three-letter code* (e.g. Ala- for Alanine).

Acknowledgment.....	iii
Publications and Poster presentations.....	iv
Abstract.....	v
Zusammenfassung.....	vii
Abbreviations.....	x
Chapter 1: Introduction.....	1
1.1 The importance of light.....	1
1.2 Photoreceptor proteins - light sensitive protein-ligand complexes that sense a variety of different wavelengths of the electromagnetic spectrum.....	2
1.2.1 Distinct classes of photoreceptor proteins.....	2
1.2.2 Blue-light photoreceptors using flavin chromophores.....	4
1.2.2.1 BLUF-proteins.....	4
1.2.2.2 Cryptochromes.....	5
1.2.2.3 LOV photoreceptors.....	7
1.2.3 Bacterial LOV photoreceptor proteins.....	8
1.2.3.1 The short LOV proteins of <i>Pseudomonas putida</i> KT2440.....	10
1.2.4 LOV signalling paradigm – the conserved photochemistry of LOV photoreceptors.....	10
1.2.5 Structure and signalling mechanism of LOV proteins - insights gained from X-ray crystallography.....	11
1.2.5.1 LOV signal-transduction involves auxillary structural elements present outside of the canonical LOV sensory core.....	13
1.2.6 Structural dynamics of LOV proteins – insights gained from solution NMR spectroscopy.....	15
1.2.7 LOV-based optogenetics.....	15
1.3 Solution NMR spectroscopy.....	16
1.3.1 The NMR phenomenon.....	18
1.3.1.1 Energy level diagram.....	19
1.3.1.2 Larmor frequency.....	19
1.3.2 Origin of the NMR signal.....	20

Table of Contents

1.3.3 Chemical shift	21
1.3.4 Higher-dimensional NMR experiments	22
1.3.5 NMR experiments for the structural study of proteins and peptides in solution	22
1.3.5.1 Structure determination of small peptides in solution.....	22
1.3.5.2 Structure determination of proteins in solution	22
1.3.5.3 NMR spectroscopy: more than a method to determine protein structures	23
1.4 Scope and outline of the thesis.....	26
Chapter 2: Materials and Methods.....	27
2.1.1 Bacterial strains	27
2.1.2 Plasmids	27
2.1.3 Oligonucleotide primers.....	28
2.2 Chemicals	29
2.3 Laboratory equipment	32
2.4 Software and Databases.....	33
2.5 Microbial media and antibiotics.....	34
2.5.1 Antibiotics.....	34
2.5.2 Full media for production of large amounts of recombinant proteins	34
2.5.2.1 Luria-Bertani (LB) medium [115]	34
2.5.2.2 Autoinduction full media (AI-FM) (modified according to Studier <i>et. al</i>) [116].....	34
2.5.3 Minimal media used for the production of isotopically labeled proteins.....	35
2.5.3.1 HD-MM (High cell density IPTG induction) minimal medium [117].....	35
2.5.3.2 Autoinduction minimal media (AI-MM medium) [117].....	36
2.5.3.2.1 Preparation of agar plates	37
2.6 Buffer and solutions	37
2.6.1 Agarose gel-electrophoresis	37
2.6.2 CaCl ₂ solution for preparation of chemically competent <i>E. coli</i> cells	38
2.6.3. Buffers for protein purification	38
2.6.3.1 Lysis Buffer	38
2.6.3.2 Ni-NTA Chromatography buffers.....	38

Table of Contents

2.6.3.3 Dialysis, NMR and protein storage buffer	39
2.6.4 Sodium-dodecyl polyacrylamide gel-electrophoresis (SDS-PAGE) [118].....	39
2.6.5 Proteolytic removal of hexa-histidine-tag.....	40
2.6.6 Buffer for HPLC-SEC experiments.....	40
2.6.7 Bradford Reagent.....	40
2.7 Enzymes	41
2.8 Peptide synthesis	41
2.9 Methods	41
2.9.1 Microbiological Methods.....	41
2.9.1.1 Cryo-preservation and long-term storage of <i>E. coli</i> strains.....	41
2.9.1.2 Growth and storage of bacterial strains.....	41
2.9.1.3 Preparation of competent cells by CaCl ₂ method [120]	42
2.9.1.4 Heat-shock transformation of chemically competent cells [120]	42
2.9.1.5 Induced gene expression in <i>E. coli</i>	42
2.9.1.5.1 Expression of PpSB1-LOV/PpSB2-LOV wild-type and truncated constructs	43
2.9.1.5.2 Expression optimization for the isotopic labeling of proteins for NMR experiments	43
2.9.1.5.3 Protein production in HD-MM (2.5.3.1) medium.....	43
2.9.1.5.4 Protein production in AI-MM (2.5.3.2) medium	43
2.9.2 DNA methods.....	44
2.9.2.1 Isolation of Plasmid DNA	44
2.9.2.1.1 Small and large-scale plasmid preparation [115].....	44
2.9.2.2 Purification of DNA from agarose gels.....	44
2.9.2.3 Agarose gel-electrophoresis [115].....	45
2.9.2.4 Recombinant DNA techniques	45
2.9.2.4.1 Restriction of DNA by restriction endonucleases [121].....	45
2.9.2.4.2 DNA ligation [122]	45
2.9.2.5 Polymerase chain reaction (PCR) [123].....	46
2.9.2.6 DNA Sequencing	46

Table of Contents

2.9.2.7 Determination of DNA concentration purity	47
2.9.3 Biochemical and protein analytical methods.....	47
2.9.3.1 Purification of heterologously expressed proteins.....	47
2.9.3.1.1 Cell lysis	47
2.9.3.1.2 IMAC chromatography on Ni-NTA resin	47
2.9.3.2 Buffer exchange by dialysis	48
2.9.3.3 Proteolytic removal of the hexa-histidine-tag by Thrombin.....	48
2.9.3.4 Concentration of protein samples by ultrafiltration	49
2.9.3.5 Determination of protein concentrations [126].....	49
2.9.3.6 Estimation of native protein molecular weights	49
2.9.3.6.1 Protein molecular weight estimation by size exclusion chromatography (SEC) using a (high-pressure) HPLC system	50
2.9.3.7 HPLC-SEC analysis of peptide samples	50
2.9.3.8 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) [118].....	51
2.9.3.9 HPLC analysis of LOV protein flavin content and flavin species distribution	52
2.9.3.10 Matrix Assisted Laser Desorption/Ionization-Time Of Flight Mass Spectrometry (MALDI-TOF MS) molecular mass analysis	53
2.9.4 Spectroscopic methods	53
2.9.4.1 Circular dichroism (CD) spectroscopy.....	53
2.9.4.2 UV/Vis spectrophotometry.....	55
2.9.4.2.1 Measurement of the LOV protein photocycle associated spectral changes	55
2.9.4.2.2 Estimation of the apo-protein content of LOV-protein preparations.....	55
2.10 Bioinformatic analysis.....	56
2.11 NMR spectroscopy	57
2.11.1 Peptide NMR measurements – sample preparation.....	57
2.11.2 LOV-protein NMR measurements – sample preparation.....	57
2.11.3 Data Collection.....	58
2.11.4 Processing	58
2.11.5 Resonance assignment for J α -peptide samples.....	58

Table of Contents

2.11.6 Sequential assignment of protein resonances	60
2.11.6.1 Backbone resonance assignment of PpSB1-LOV	60
2.11.7 Chemical shift perturbation	61
2.11.8 NMR structure calculation	62
2.11.9 Hydrogen-deuterium (H/D) exchange	62
2.11.10 Heteronuclear ¹ H- ¹⁵ N NOE (hetNOE), longitudinal and transverse relaxation experiments	63
2.11.11 Deposition of NMR assignments	63
Chapter 3: Results	65
3.1 Conservation of auxiliary structural features in the two <i>P. putida</i> KT2240 short LOV proteins	66
3.1.1 Cloning and expression of PpSB1-LOV and PpSB2-LOV truncated variants	67
3.1.2 Bioinformatic analysis of the N- and C-terminal auxiliary sequence elements of PpSB1- LOV and PpSB2-LOV	69
3.1.3 Structural studies on the isolated PpSB1-LOV and PpSB2-LOV J α -peptides	73
3.1.3.1 Quality of the custom synthesized PpSB1-J α and PpSB2-J α peptides	73
3.1.3.2 CD spectroscopic studies reveal independent folding of the isolated PpSB1-J α and PpSB2-J α peptides	74
3.1.3.3 Addition of the helix-stabilizing agent 2,2,2-trifluoroethanol (TFE) is necessary to stabilize the two peptides for NMR studies	77
3.1.3.4 NMR experiments	79
3.1.3.5 Structure determination of PpSB1/PpSB2-J α peptide by NMR	79
3.2 Solution NMR experiments using the full-length PpSB1-LOV protein	83
3.2.1 Optimization of expression and purification of PpSB1-LOV for NMR spectroscopic studies	83
3.2.1.1 Optimization of expression and purification conditions – Expression in <i>E. coli</i> BL21 (DE3)	83
3.2.1.1.1 Chromophore-load of the <i>E. coli</i> BL21 prepared PpSB1-LOV protein analyzed by UV/Vis spectrophotometry	84
3.2.1.1.2 Chromophore-composition of the <i>E. coli</i> BL21 prepared PpSB1-LOV protein analyzed by HPLC	85
3.2.1.2 Optimization of expression and purification conditions - Expression in <i>E. coli</i> CmpX131 (DE3)	86

Table of Contents

3.2.1.2.1 Chromophore-load and -composition of the from <i>E. coli</i> CmpX131 prepared PpSB1-LOV protein	88
3.2.2 Initial sample preparation for NMR studies	89
3.2.2.1 Blue-light sensitivity of the isotopically labeled PpSB1-LOV sample	90
3.2.2.2 Matrix assisted laser desorption/ionization mass spectrometry (MALDI-TOF MS).....	91
3.2.2.3 Optimization of buffer/measuring conditions for NMR experiments	92
3.2.2.4 Final sample preparation procedure used for all NMR experiments	94
3.2.2.5 Final dark-state ^1H - ^{15}N HSQC spectra	96
3.2.2.6 A custom made illumination device allows sample illumination inside the NMR spectrometer	96
3.2.2.7 Comparison of dark- and light-state PpSB1-LOV HSQC spectra suggest widespread structural changes upon illumination of the protein	97
3.2.2.8 Sample integrity under NMR conditions.....	99
3.2.2.8.1 Photosensitivity and dark recovery under NMR conditions	99
3.2.2.8.2 Determination of the native quaternary structure of PpSB1-LOV under NMR measurement conditions.....	100
3.2.3 Backbone resonance assignment of the PpSB1-LOV dark- and light-state	102
3.2.3.1 Backbone assignment for the PpSB1-LOV dark-state	104
3.2.3.2 Backbone and side-chain assignment for the light-state	105
3.2.4 PpSB1-LOV secondary structure information obtained from NMR data	108
3.2.5 PpSB1-LOV structural dynamics studied by NMR.....	110
3.2.5.1 Dynamics of PpSB1-LOV in the dark-state	110
3.2.5.2 Dark and light state solvent accessibility studied by proton-deuterium (H/D) exchange experiment	112
Chapter 4: Discussion	116
4.1. The PpSB1-LOV light-state X-ray structure probed in solution.....	116
4.1.1 Structural features of the recently solved PpSB1-LOV X-ray structure.....	116
4.1.2 Probing the role of the N- and C-terminal extensions for the structural integrity of the <i>P. putida</i> LOV proteins.....	118
4.1.2.1 PpSB1-LOV appears to be dimeric in both the dark- and light-state.....	118

Table of Contents

4.1.2.2 The helical N- and C-terminal extensions of PpSB1-LOV and PpSB2-LOV are essential for folding and structural integrity of both proteins	118
4.1.2.3 The isolated C-terminal extension of PpSB1-LOV and PpSB2-LOV adopt their native secondary structure independent from the rest of the protein	119
4.2 Light-state NMR studies hint at a grossly similar structure of PpSB1-LOV in solution and the crystal	122
4.2.1 Comparison of NMR-based secondary structure predictions for PpSB1-LOV in the light-state and the crystal structure	122
4.2.2 H/D exchange and relaxation studies support the $J\alpha$ -helix orientation observed in the light-state X-ray structure of PpSB1-LOV	125
4.3 On the nature of PpSB1-LOV photo-activation	126
4.3.1 Chemical shift perturbation analyses suggest a potential mode for LOV-photo-activation	126
4.3.2 Further NMR spectroscopic evidence for the proposed intra-molecular signal-relay mechanism	132
4.3.2.1 NMR-based secondary structure predictions for the dark- and light-state of PpSB1-LOV suggest local secondary structural changes in A β -B β region	132
4.3.2.2 Relaxation experiments	137
4.4 Implications of the here presented NMR data for the conservation of the LOV photo-activation and signal-relay paradigm	138
4.4.1 Photo-activation of the plant phototropin LOV2 domain involves dissociation of the $J\alpha$ -linker helix	140
4.4.2 Light-induced transient dimerization of the short LOV protein VVD triggers protein-protein interactions in <i>N. crassa</i>	141
4.4.3 Light-induced DNA binding by an LOV-HTH protein is mediated by sensor / effector domain rearrangement	141
4.4.4 Light-induced rotation of coiled-coil forming $J\alpha$ -linker helices in the dimeric YtvA photoreceptor from <i>B. subtilis</i>	142
4.5 Conclusions	143
4.6 Outlook	145
References	146
Appendix A	155

Chapter 1: Introduction

1.1 The importance of light

“Every major world religion speaks a language of light.

On Christmas, billions of believers say Christ was born to light the world, fulfilling God's first gift to the heavens and earth: "Let there be light."

On Diwali, billions of Hindus, Sikhs and Jains light lanterns and small earthen lamps to make nights glow. Lights awaken awareness of God and the triumph of good over evil.

Muslims speak of Allah as the source of light - inspiring, motivating and guiding God's people” [1].

These are a few selected examples that describe and account for the importance of light for billions of believers on our planet. Yet, whether we are religious or atheist for all of us humans light remains the most important visual stimulus. Likewise, plants and phototrophic microbes depend on it as a source of energy generated in the process of photosynthesis [2, 3]. Thus, light represents one of the most important environmental cues for most living beings. Nonetheless, the physical nature of light remained for a long time largely elusive. Variable scientific descriptions of the phenomenon light were widespread throughout the scholarly world since ancient Greek and Roman times [4, 5].

In modern physics, light is thought to be a form of electromagnetic radiation whose physical properties can be explained by describing it either as a particle (photon) or wave [6]. The human eye can only perceive the visible fraction of the electromagnetic spectrum (approx. 380 nm - 740 nm) [7]. Visible light can be utilized as source of energy by many kinds of organisms whereas shorter wavelength light (e.g. ultraviolet light < 380 nm) can cause considerable harm, i.e. by acting on biological macromolecules such as DNA or protein-bound ligands [8]. Therefore the correct perception and integration of environmental light signals is essential for the survival and adaptation of most organisms. In more indirect terms light can also function as an information carrier, allowing us, other animals as well as plants and microbes to respond to changing day and night, as well as seasonal cycles [9].

1.2 Photoreceptor proteins - light sensitive protein-ligand complexes that sense a variety of different wavelengths of the electromagnetic spectrum

At the molecular level, changes in external environmental light conditions are frequently detected by so-called sensory photoreceptor proteins [10]. Hereby, the protein moiety as a folded polypeptide chain composed of the 20 proteinogenic amino acids, cannot absorb the visible fraction of the electromagnetic spectrum directly. For photon absorption photoreceptors thus usually rely on a co-factor that is bound within the respective protein [10]. As the light-absorbing component of the photoreceptor protein these co-factors are called chromophores derived from the Greek word for color (*chromos*). After light absorption an electronic excited state of the chromophore is populated, representing basically a reorganization of the electron distribution of the molecule. The chromophore can only absorb a certain light quality when the energy difference between two different molecular orbitals of the chromophore molecule falls within the range of the energy content of the respective radiation. Thus, depending on the different chromophore species bound within the various photoreceptor proteins, different absorptive characteristics are achieved enabling the response to different wavelengths of the visible spectrum. Depopulation of the excited state in turn can occur via radiative (fluorescence) and non-radiative processes or facilitates subsequent photochemical reactions which enable receptor photocycling (see below). For all photoreceptor proteins, illumination usually causes formation of a meta-stable signaling state, which is assumed to possess different structural properties than the respective protein in the dark. This structural change in turn is assumed to relay the light signal within the cell. Possible mechanisms of this signal relay, propagation and amplification will be outlined below.

1.2.1 Distinct classes of photoreceptor proteins

According to the bound chromophore, photochemistry and detected wavelength fraction, photoreceptor proteins with a wide distribution in the three domains of life (bacteria, archaea and eukaryotes) can be divided into eight distinct classes: the rhodopsins [11, 12], the phytochromes [13], xanthopsins [14], the orange carotenoid proteins (OCP)[15], plant UV photoreceptors such as UVR8 [16], sensors of blue light using flavin-adenine dinucleotide (FAD) family (BLUF-proteins) [17], cryptochromes [18] and light, oxygen, voltage (LOV) proteins [19, 20] (Fig. 1).

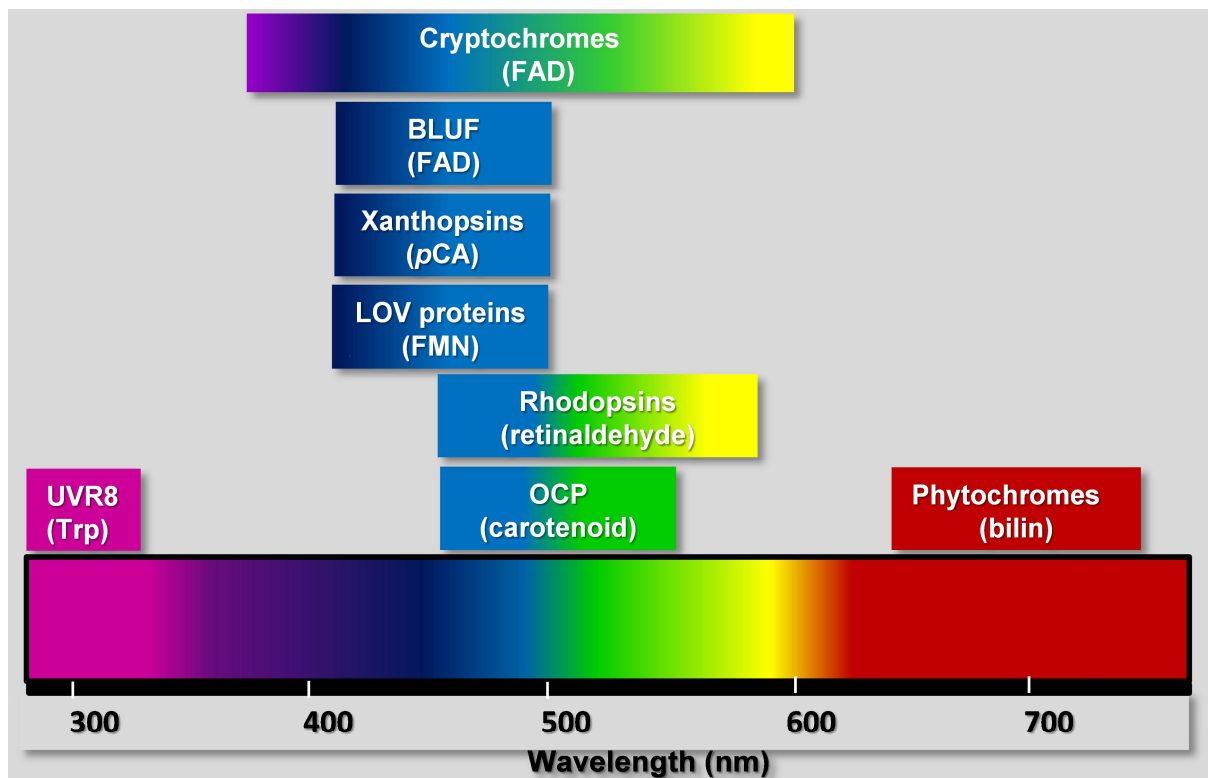


Figure 1: Photoreceptors proteins with a wide distribution in the three domains of life. The different classes of photoreceptors, their absorptive characteristics (absorbed wavelength fraction) as well as the bound chromophore species (in brackets) are shown above the corresponding fraction of the UV/Vis spectrum. Abbreviations: BLUF: sensor of the blue light using FAD, LOV: light, oxygen voltage, OCP: orange carotenoid protein, FAD: flavin-adenine-dinucleotide, FMN: flavin mononucleotide, pCA: *para*-coumaric acid, Trp: tryptophane.

For the first three families, light absorption causes *cis-trans* isomerization of a double bond within the chromophore [12-14]. The last three families are using flavin derivatives as chromophore but employ a different photochemistry for photocycling. The biological function of most photoreceptor classes including signal transduction mechanisms as well as their photocycle after illumination have been studied widely [21-23]. However, detailed aspects regarding photocycle mechanistic questions, structural consequences of signalling state formation as well as downstream regulatory networks remain elusive in many cases. Reviewing all those aspects for the known eight photoreceptor families would be outside the scope of this introduction. Therefore, only the most important concepts, i.e. with relevance to the here presented work, will be summarized in the following chapters.

1.2.2 Blue-light photoreceptors using flavin chromophores

BLUF-domain containing proteins, cryptochromes (CRYs) and all LOV photoreceptors carry blue-light absorbing flavin derivatives as chromophore [24]. Hereby most LOV photoreceptor proteins, bind FMN while the BLUF and CRY proteins bind FAD as chromophore [17, 25]. One exception from the rule is the *Neurospora crassa* VVD protein, which is a LOV photoreceptor but binds FAD instead of FMN [26]. While LOVs, BLUFs and CRYs contain very similar light absorbing chromophoric structures, i.e. represented by the flavin isoalloxazine ring, their photochemistry as well as their biological function and signal-transduction mechanisms initiated by blue-light absorption are different as will be discussed in the following:

1.2.2.1 BLUF-proteins

The acronym BLUF stands for sensors of Blue-Light Using FAD [17]. In most BLUF photoreceptors one BLUF sensor domain is attached N-terminally to a so-called effector domain. Hereby absorption of light by the BLUF domain modifies the activity of the attached effector [27]. The BLUF photocycle apparently involves a light-dependent rearrangement of the hydrogen-bonding network surrounding the FAD isoalloxazine ring, resulting in a spectrally red-shifted signalling state ($AppA_{sig}$) (Fig. 2, B) [28]. Hereby, two amino acid residues (Tyr21 and Gln63, *Rhodobacter sphaeroides* AppA-BLUF numbering) were found to be essential for BLUF photocycling [29] (Fig. 2, A). The BLUF domain was first discovered as the N-terminal part of the flavoprotein AppA from *R. sphaeroides* (Fig. 2, C) [17, 30]. AppA acts as a transcriptional regulator of photosynthesis gene expression in the phototrophic organism [30], [17]. Since the discovery of AppA, BLUF proteins have been identified in *E. coli* [31], the unicellular flagellate *Euglena gracilis* [32], as well as in cyanobacteria such as *Synechocystis* sp. PCC6803 [33] and *Thermosynechococcus elongates* BP-1 [34]. Physiological responses mediated by BLUF photoreceptors include regulation of gene expression (*E. coli*) and negative phototaxis (*E. gracilis* and *Synechocystis*) [32, 35].

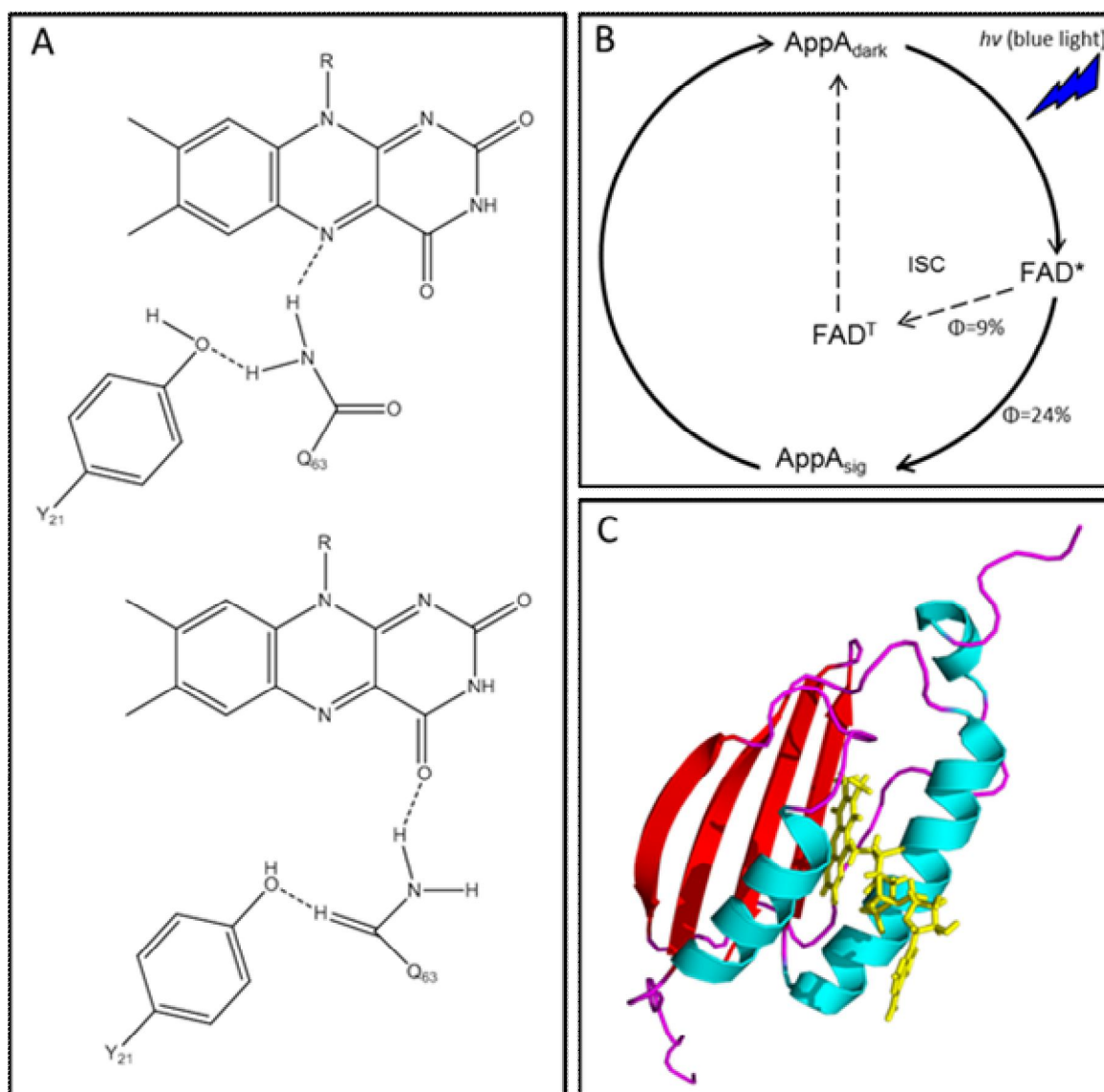


Figure 2: (A) The residues corresponding to Tyr21 and Gln63 of AppA are essential amino acids involved in BLUF photocycling. Gln63 is hydrogen bonded to the hydroxyl group of Tyr21. In different crystal structures different rotamers of this residue were found: (top): Q63_{N_e}-Y21 and (bottom): Q63_{N_e}-FAD-C(4)=O [36]. Upon blue-light illumination the hydrogen-bonding network surrounding the FAD chromophore is rearranged resulting in a by 10 nm spectrally red-shifted signalling state (AppA_{sig}). R= C₁₅H₂₅N₅O₁₃P₂ (B) Photocycle scheme for the AppA BLUF domain [28]. (Reprinted with permission from [28]. Copyright (2005) American Chemical Society.) (C) X-ray structure of the dark-state of the AppA BLUF domain from *R. sphaeroides* (pdb entry: 2IYG) [37]. The FAD chromophore is shown in stick representation in yellow. α -helices are shown in cyan, β -sheets are depicted in red and random coils are highlighted in pink.

1.2.2.2 Cryptochromes

Cryptochrome (CRY) photoreceptors were first identified in plants [38]. The name was initially derived from the Greek term (*krypto chroma*, meaning hidden color). Homologous

CRY proteins have later been identified in lower and higher eukaryotes as well as in prokaryotes [39]. In the ground state, plant CRYs such as *Arabidopsis thaliana* CRY1, bind fully oxidized FAD which is photoreduced to the corresponding semireduced semiquinone (neutral radical FADH[•]) upon blue-light illumination.

The semiquinone form of FAD in turn can absorb blue/green light to become fully reduced (FADH⁻) [40, 41] (Fig. 3, A, B). With regard to sequence similarity, cryptochromes can be divided in three families: plant cryptochromes, animal cryptochromes and the so-called cryptochrome-DASH proteins. Plant cryptochromes (such as *A. thaliana* CRY1, Fig. 3, C) regulate hypocotyl elongation, seed germination and pigment accumulation and phototropism [42, 43]. Animal cryptochromes have been found in insects, fish, amphibians and mammals acting as components of the circadian clock machinery [44, 45]. The regulatory role of the CRY-DASH proteins is still unknown [25]. Hereby the acronym DASH stands for the organisms that contain homologous proteins: *Drosophila*, *Arabidopsis*, *Synechocystis* and Human. Recent hypothesis and studies imply CRYs as the magnetoreceptor of birds and butterflies involved in magnetic orientation during migration [46].

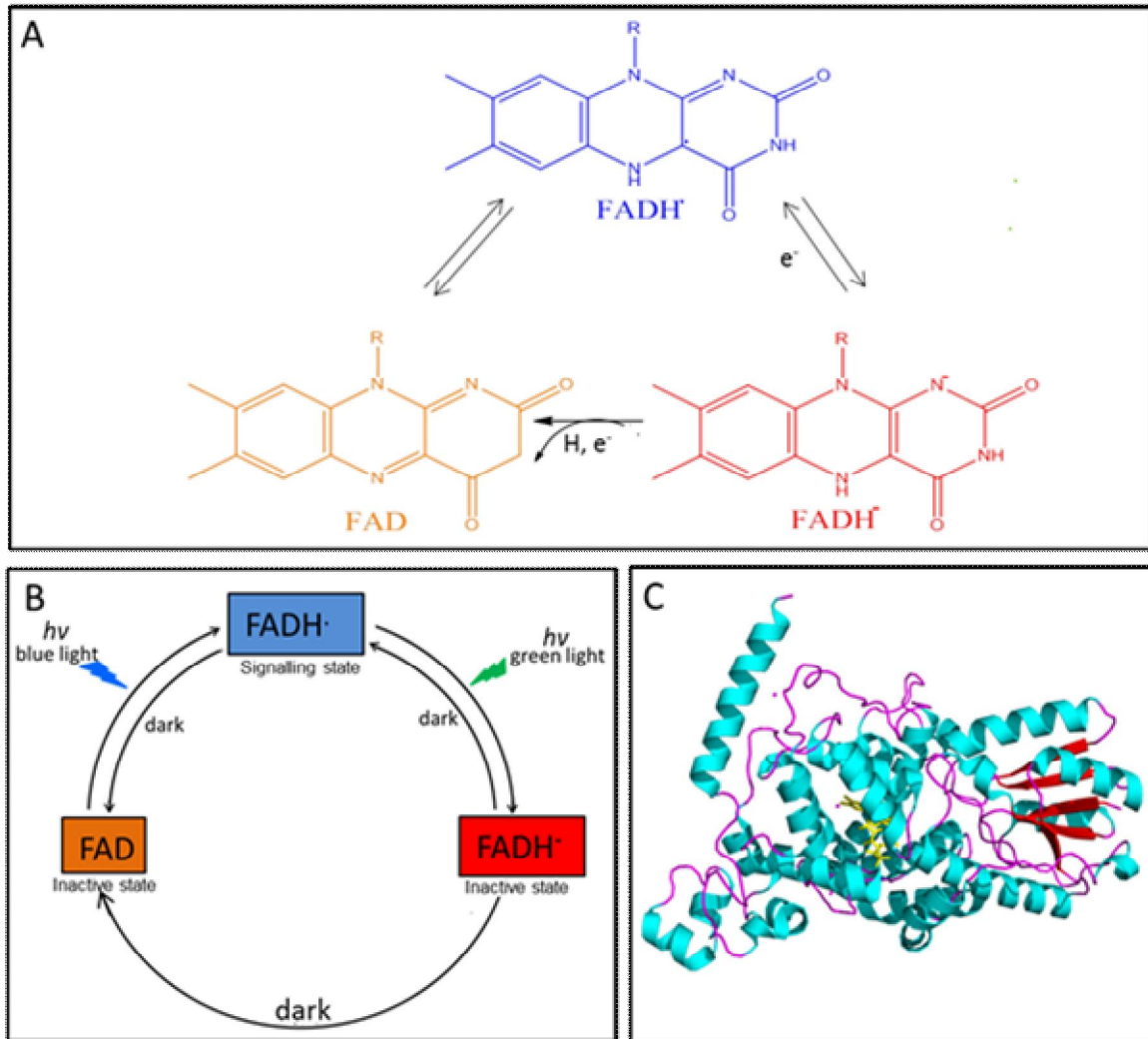








Figure 3: (A) The three possible redox forms of flavins found in CRY's are shown, oxidized (FAD, orange), semireduced semiquinone (FADH[•], blue) and fully reduced hydroquinone FADH⁻, red) [40], R= C₁₅H₂₅N₅O₁₃P₂ (B) A schematic representation of the photocycle of cryptochrome [41] (C) X-ray structure of CRY1 from *A. thaliana* in the dark-state (pdb entry: 1U3C). The FAD chromophore is depicted in stick representation in yellow. α -helices are shown in cyan, β -sheets are depicted in red and random coils are highlighted in pink.

1.2.2.3 LOV photoreceptors

LOV photoreceptors represent one of the most versatile blue-light photoreceptor families known to date. LOV domains, the sensory modules of LOV photoreceptors, are widespread in the three kingdom of life [42]. In eukaryotes six distinct LOV photoreceptor architectures can be distinguished (Table 1): plant and algal phototropins, fern neochromes, algal aureochromes, fungal WC-1 and VVD proteins as well as the recently described LOV

histidine kinases (LOV-HK) identified in picoalgae such as *Ostreococcus tauri* (Table 1) [23, 42, 43].

Table 1: Schematic representation of the six major eukaryotic LOV photoreceptor families, including organismal distribution, multi-domain architecture and physiological functions.

Family	Distribution	Architecture	Functions
Phototropin	Green plants: <i>A. thaliana</i> <i>Avena sativa</i> <i>Chlamydomonas reinhardtii</i>		phototropism, phototaxis, chloroplast relocation and stomatal-opening
Neochrome	Ferns: <i>Adiantum capillus-veneris</i>		phototropism and chloroplast relocation
Aureochrom e	Algae: <i>Vaucheria frigida</i>		Photomorphogenesis
WC- 1	Fungi: <i>N. crassa</i>		setting of circadian clock
VVD	Fungi: <i>N. crassa</i>		setting of circadian clock
LOV-HK	Picoalgae: <i>O. tauri</i>		setting of circadian clock






Abbreviations: LOV: light, oxygen, voltage; bZIP: leucin zipper helix-loop-helix domain, PAS: Per-Arnt-Sim domain, ZnF: zinc-finger DNA binding domain, Kinase: serine/threonine kinase (phototropin, neochrome), histidine kinase (LOV-HK), REC: cheY-homologous receiver domain

1.2.3 Bacterial LOV photoreceptor proteins

Despite its wide distribution in eukaryotes, genome mining revealed the widespread occurrence of putative LOV photoreceptors in phototrophic and chemotrophic prokaryotes [44]. As of late 2011 over 452 mostly putative bacterial LOV proteins have been identified in bacteria (Homepage Dr. Aba Losi, University of Parma, Italy) (<http://www.fis.unipr.it/~losia/Lov-proteinstable.pdf>). For example LOV domain containing photoreceptors have been found in a variety of plant and animal pathogens such as

Pseudomonas syringae [45], *Brucella abortus* [46], in marine microbes such as *Caulobacter crescentus* [47], in some plant root colonizing species like *Pseudomonas putida* [48, 49] as well as in common soil bacteria like *Bacillus subtilis* [50-52]. While, all so far described prokaryotic LOV photoreceptors possess a classical phototropin LOV-like photochemistry (see below), until recently only little was known about their physiological function.

Table 2: Exemplary bacterial LOV photoreceptor multi-domain architectures [55]. The LOV sensor domain (in blue) is usually N-terminally attached to different effector domains. The highlighted short LOV proteins of *Pseudomonas putida* KT2440 (blue underlined) and other *Pseudomonas* species do not possess any fused effector domain.

Family	Distribution	Architecture	Functions
PST-LOV	Plant pathogens: <i>P. syringae</i>		unknown
LOV-HK	Animal pathogens: <i>B. abortus</i>		Virulence regulation
LovK	Marine microbes: <i>C. crescentus</i>		Cell adhesion
PpSB1-LOV / PpSB2-LOV	Plant root colonizing: <i><u>P. putida</u></i>		unknown
YtvA	Soil bacteria: <i>B. subtilis</i>		Stress response

Abbreviations: RR-response regulator; PAS- Per-Arnt-Sim; HK-histidine kinase; STAS-sulphate transporter anti-sigma factor antagonist domain.

As in the eukaryotic world, bacterial LOV photoreceptor proteins are usually multi-domain sensory systems consisting of a single LOV sensor domain, associated with different effector domains, such as kinases, helix-turn-helix DNA-binding domains, anti-sigma factors, transcriptional regulators, phosphodiesterases and cyclases [53] (Table 2). Some bacterial LOV proteins as well as fungal LOV proteins exist that lack any distinguishable fused effector domain (Table 1 and 2). Examples include the LOV proteins from *P. putida* KT2440, the VVD protein from *N. crassa* and the RsLOV protein from *R. sphaeroides* [48, 54, 55]. In

bacteria these kind of proteins are referred to as short LOV proteins [48]. They constitute with about 11-13% of all bacterial LOV proteins, the third largest bacterial LOV photoreceptor family [53].

1.2.3.1 The short LOV proteins of *Pseudomonas putida* KT2440

The genome of *P. putida* KT2440 contains two genes encoding for two short LOV proteins PpSB1-LOV (Q88E39) and PpSB2-LOV (Q88JB0). Both yet putative LOV photoreceptors consist of a conserved LOV core flanked by N- and C- terminal extensions but lack a fused effector domain. The physiological role of the two LOV proteins as well as the importance of blue light for the saprotrophic plant-root associated γ -Proteobacterium remains so far elusive. However, yet unpublished experiments performed in our laboratory hint at the involvement of PpSB1-LOV in the iron-starvation response of the organism (Katrin Jentzsch, Ph.D. thesis 2012, unpublished).

1.2.4 LOV signalling paradigm – the conserved photochemistry of LOV photoreceptors

Common to all to date characterized LOV photoreceptor proteins is their photochemistry. Upon blue-light illumination, LOV domains undergo a spectroscopically but not structurally well characterized photocycle. Based on steady-state and time resolved spectroscopic as well as X-ray crystallographic data the following LOV photocycle intermediates have been identified (Fig. 4, A, B): In the dark-state, FMN is non-covalently bound within the protein displaying a typical flavin absorption maximum at 447 nm (LOV₄₄₇). Upon blue-light absorption by the chromophore, the excited singlet state of FMN is populated within picoseconds [56]. The excited singlet state decays via intersystem crossing to the excited triplet state [57]. The triplet excited state, also denoted as LOV₆₆₀, absorbs maximally at around 660 nm and decays within a few microseconds to give rise to a covalent adduct between a totally conserved cysteine residue in the LOV domain (C450 in *A. sativa* phot1-LOV2, C62 in *B. subtilis* YtvA and C53 in PpSB1/2-LOV) and the C4a carbon atom of the FMN isoalloxazine ring [58-61]. As the longest lived intermediate this photoadduct represents the light- or signaling state of LOV photoreceptor proteins, possessing a characteristic absorption maximum at around 390 nm [62]. This state is hence often referred to as LOV₃₉₀. From the signaling state the LOV protein returns in dark to the parent-, ground- or dark-state (LOV₄₄₇) within seconds to hours depending on the protein [49, 63, 64]. For example, the

two *P. putida* LOV proteins that share about 66 % identical amino acid positions, display drastically different photocycle kinetic properties. At 25 °C PpSB1-LOV reverts back to the dark-state in about 40 hours, whereas PpSB2-LOV possesses a dark recovery time constant of about 140 s [48, 49].

Like for all other photoreceptors, signaling state formation is assumed to be accompanied by a conformational rearrangement in the surrounding protein that is propagated to the surface of the sensor domain.

1.2.5 Structure and signalling mechanism of LOV proteins - insights gained from X-ray crystallography

Initial X-ray crystallographic studies performed using isolated LOV sensor domains of plant and algal phototropins revealed a typical Per-Arnt-Sim (PAS) fold for the LOV sensory core domain [65-67]. The name PAS was hereby derived from the names of the proteins for that it was first described, the *Drosophila* protein period (Per), the aryl hydrocarbon receptor nuclear translocator protein (ARNT) and the *Drosophila* single-minded protein (Sim) [68]. The PAS fold consist of five antiparallel β -sheets (A β , B β , G β , H β and I β) surrounded by four α -helices (C α , D α , E α and F α) that together form a tight pocket mediating ligand (FMN in case of LOV domains) binding [66]. The topological order of secondary structure elements typical for LOV domains is A β -B β -C α -D α -E α -F α -G β -H β -I β (Fig. 4, C) [69]. A conserved sequence motif GXN**C**RFLQ is found on helix E α and contains the totally conserved covalently binding cysteine (bold, underlined).

Initially LOV domain X-ray structures were determined for the phot1-LOV1 domain from the green alga *C. reinhardtii* [70], the LOV2 domain of the neochrome protein (Phy3) of *A. capillus-veneris* [71], the phot1-LOV1 and the phot2-LOV1 domains from *A. thaliana*, the phot1-LOV2 domain of *A. sativa* as well as for the *N. crassa* VVD protein [54]. Usually, LOV protein crystals are grown in the dark. Diffraction data are collected to obtain the dark-state structure of the protein. Subsequently crystals are illuminated with blue light to obtain light state structural information [54, 66, 70, 72]. While photoadduct formation can be readily observed in illuminated LOV protein crystals, large scale conformational changes could not be detected when dark grown crystals were illuminated. Nevertheless, in the immediate vicinity of the FMN chromophore minor structural changes could be observed. One common

theme usually discussed in terms of LOV photo-activation is a light-induced flipping of the side chain of a totally conserved glutamine residue that is part of the hydrogen-bonding network stabilizing FMN binding in the LOV domain. In *B. subtilis* YtvA-LOV, residue Q123 and Q513 in *A. sativa* phot1-LOV2, the respective glutamine forms a hydrogen bond with the O4 atom of the FMN ring. Upon illumination, the glutamine undergoes a conformational rearrangement, probably involving a flip of its side-chain to form a new hydrogen bond with the newly protonated N5 atom of the FMN ring [70, 72, 73].

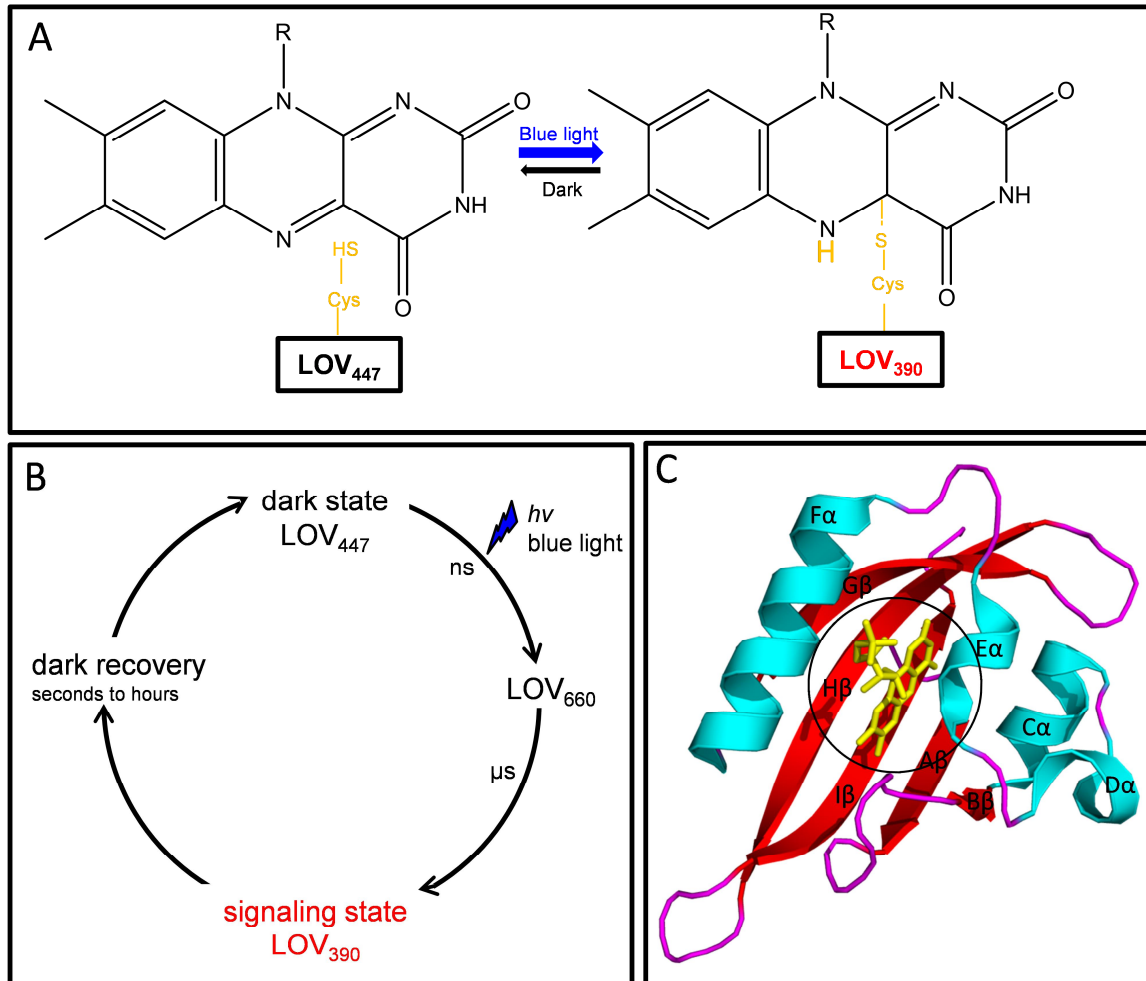


Figure 4: (A) Primary photoreaction of LOV domains. FMN is non-covalently bound to the protein in the dark state showing an absorption maximum at 447 nm (LOV₄₄₇). Upon blue-light absorption by the chromophore a covalent bond is formed between the conserved cysteine and C4a carbon atom of the FMN isoalloxazine ring. The light state displays an absorption maximum at 390 nm (shown in red color) [64], R= C₅H₁₃O₇P (B) Photocycle of LOV proteins including time regime of the respective decay/interconversion processes (C) X-ray structure of the LOV domain from *C. reinhardtii* in the dark state (pdb entry: 1N9L). The FMN chromophore is shown in stick representation (yellow). α -helices are shown in cyan, β -sheets are depicted in red and random coils are highlighted in pink. All structural elements are labelled in topological order from A β to I β .

Interpolated from this and other minor structural changes observed upon illumination of dark grown LOV protein crystals, several hypothesis have been brought forward regarding intra- and intermolecular signal-relay mechanisms.

1.2.5.1 LOV signal-transduction involves auxillary structural elements present outside of the canonical LOV sensory core

X-ray crystallographic data for an extended *A. sativa* phot1-LOV2 construct suggested the involvement of N- and C-terminally located auxiliary helical structures in the light-dependent signal relay. In dark grown crystals the N-terminal helical cap ($A'\alpha$) as well as the C-terminal $J\alpha$ -helix interact with the β -sheet scaffold, burying hydrophobic residues of the core domain from the solvent [70] (Fig. 5, A). Upon illumination flavin-C4a-cysteinyl thiol adduct formation as well as other minor structural changes in the vicinity of the flavin chromophore are observed [70]. Based on this data, the authors suggested that the structural changes in the FMN vicinity are relayed to the surface of the protein namely to the N-terminal and C-terminal helical structures via a network of hydrogen-bonding residues, located on $I\beta$, $A\beta$ and the $A'\alpha$ - $A\beta$ loop [70].

For the short LOV protein VVD from *N. crassa* it has been shown that photo-activation induces structural changes near a partially helical N-terminal cap (helix $A'\alpha$) which in the full-length short LOV protein covers the central β -sheet surface of one LOV subunit and makes contact with the N-terminal cap of the opposite subunit in the dimer (Fig. 5, B). As for AsLOV2, this helical cap is located outside of the conserved LOV domain [71]. With respect to signal-relay in VVD, it was suggested that light absorption triggers dimerization of the short LOV protein, which is initiated by the aforementioned structural rearrangement of the $A'\alpha$ dimer interface [72].

The first described bacterial LOV protein was the LOV-STAS photoreceptor YtvA from *B. subtilis* [73]. At the same time the YtvA LOV domain represents the first bacterial LOV protein structure obtained by X-ray crystallography. The crystallized YtvA-LOV construct consisted of a conserved sensor LOV domain, contained a short C-terminal extension constituting the interdomain linker between the LOV and STAS domains ($J\alpha$), but lacked the first 19 N-terminal amino acids [73]. In both light and dark-state the asymmetric unit contains two subunits which form a parallel head-to-head dimer. The YtvA-LOV dimer structure is

mainly stabilized by intersubunit interactions between the β -sheet scaffold of the two subunits and the C-terminally protruding $J\alpha$ -helices [73] (Fig. 5, C). While large scale light-dependent conformational changes were not detectable in YtvA-LOV, the authors suggested a minute rotation (5°) of the two monomers as a result of illumination. This in turn would result in a minor displacement of the C-terminal STAS effector domain present in the full-length protein which would have to enable the physiological response [73].

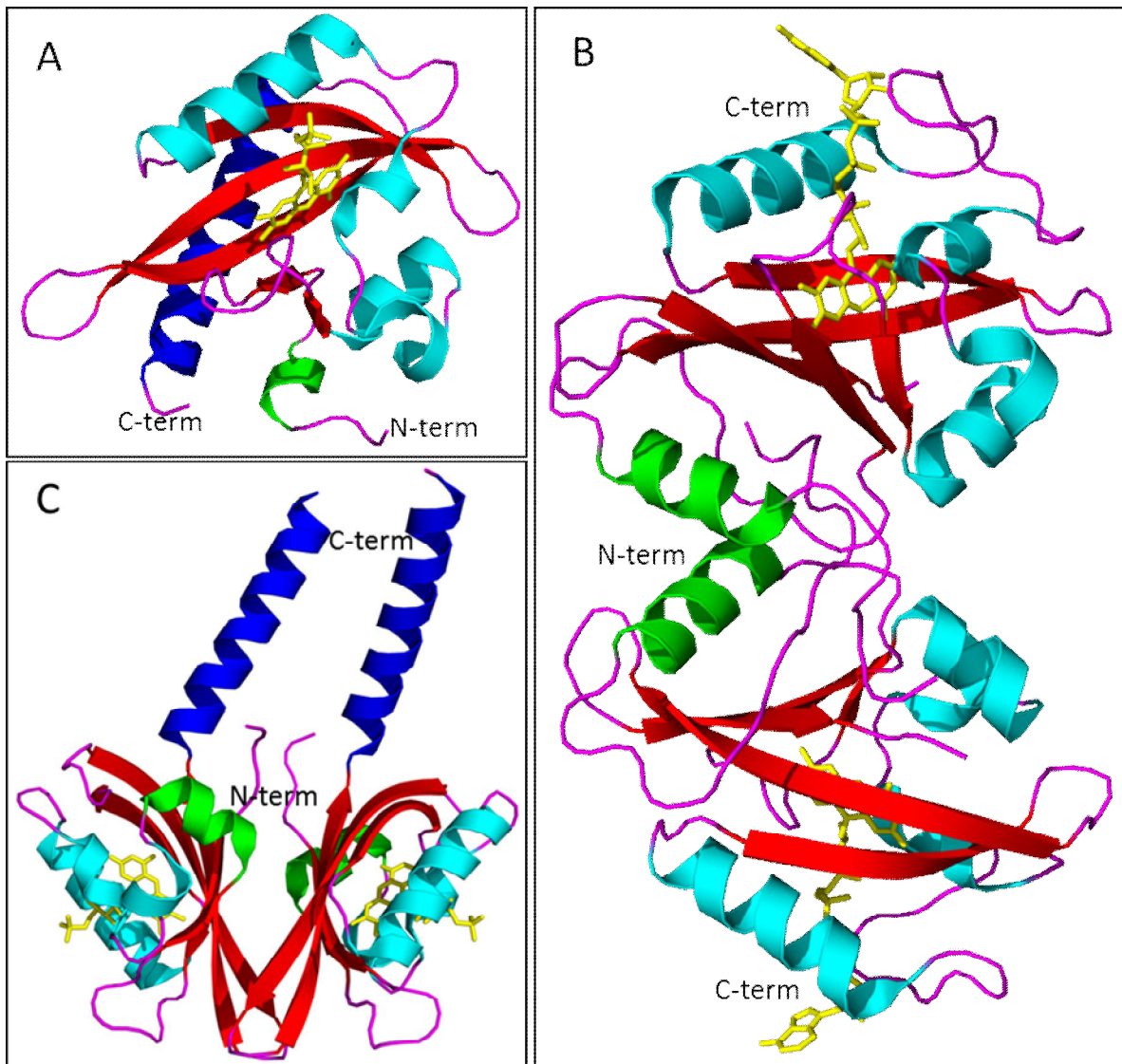


Figure 5: (A) Dark-state crystal structure of the LOV2 domain of phot1 from *A. sativa* (pdb entry: 2V1A) [70] (B) X-ray structure of the short LOV protein VVD from *N. crassa* in its dark state (pdb entry: 3RH8) [72] (C) Dark-state structure of the *B. subtilis* YtvA-LOV domain (pdb entry: 2PR5) [73]. In the LOV core domain α -helices are shown in cyan, β sheets in red and the FMN chromophore in yellow (sticks). The N-terminal A' α -helices are highlighted in green and C-terminal $J\alpha$ -helices are shown in blue.

1.2.6 Structural dynamics of LOV proteins – insights gained from solution

NMR spectroscopy

The first study that reported large scale light-induced conformational changes in a LOV sensor domain used solution NMR techniques to study LOV photo-activation and intramolecular signal relay. In this study the authors used a C-terminally extended *A. sativa* phot1-LOV2 domain construct (extended by 40 amino acids). The authors could show that in the dark the C-terminal extension forms an amphipathic α -helix (named $J\alpha$) docked against the LOV domain β -scaffold. ^1H - ^{15}N Heteronuclear Single Quantum Coherence (HSQC) spectra were recorded for the light- and dark- state of the protein revealing pronounced light-dependent minimum chemical shift changes throughout the whole protein. Based on the disappearance of a set of Nuclear Overhauser Effect (NOE) cross peaks it was suggested that the $J\alpha$ -helix dissociates from the LOV2 core and most probably becomes unstructured upon illumination [71]. Moreover, it was shown that $J\alpha$ -helix undocking mediates phototropin kinase activation in the full-length photoreceptor [74]. Similarly to previous hypotheses based on X-ray data [75] recent NMR work on *A. sativa* phot1-LOV2 revealed that the largely ignored N-terminal helix A' α (N-cap) plays an important role in initiating the light-dependent conformational change in AsLOV2. This result was supported using large scale mutational analysis, CD and NMR data as well as molecular dynamics (MD) simulations [76]. It was shown that the absence of the N-terminal A' α helix results in the constitutive undocking of the C-terminal $J\alpha$ -helix thus hinting at an interaction between the two structural elements in the dark. This interaction, as well as the packing of the two auxiliary helices against the LOV-core domain is apparently perturbed by blue-light illumination [76].

1.2.7 LOV-based optogenetics

The described modular nature as well as the reversibility (photocycling) of LOV photosensory systems has recently been exploited to engineer LOV-domain containing artificial photoreceptors which have been used as genetically encoded photoswitches allowing spatiotemporal control of biological functions *in vitro* and *in vivo* [77-83]. This relatively new line of research is part of the rapidly growing field of optogenetics, which uses optical methods (*opto-*) for stimulation and control of biological processes employing genetically encoded (photoreceptor) systems (*-genetics*) [84]. Examples for such artificially designed LOV-based optogenetic systems constructed include the photoactivated Trp-repressor protein

(LovTAP) [79], a photoactivatable LOV-GTPase fusion protein (PaRac1) [80], engineered LOV-histidine kinases [81], tunable light-inducible dimerization tags (TULIPs) based on a synthetic interaction between the LOV2 domain of *A. sativa* phot1 and an engineered PDZ domains [77] and other short peptides [82], the design of a photo-switchable dihydrofolate reductase (PAS-DHFR) [78] as well as the design of a photo-switchable LOV-dependent restriction endonuclease [83]. In conclusion, both for the understanding of LOV-dependent biological processes as well as for the structure-guided engineering of artificial photoreceptors for optogenetic purposes a detailed structural understanding of the light-induced photochemistry and the resulting protein-structural changes is indispensable.

1.3 Solution NMR spectroscopy

Given the dynamic nature of the LOV sensory system, NMR techniques appear better suited for the study of LOV domain photo-activation and signal-relay mechanisms compared to X-ray crystallography. In the presented thesis a variety of NMR techniques was used to study, PpSB1-LOV, a full-length short LOV protein of *P. putida* KT2240 [49]. Therefore, the underlying physical principles as well as standard experiments of modern NMR spectroscopic methods will be outlined in the following paragraphs.

The nuclear magnetic resonance (NMR) phenomenon was discovered by Purcel *et.al* in 1946 [85], when he detected weak radiofrequency signals generated by the nuclei of atoms in ordinary matter. At the same time Bloch *et.al* [86] independently performed different experiments and detected radiofrequency signals emitted from the atomic nuclei in water. They observed that magnetic nuclei could absorb the energy of radiofrequency pulses when placed in a magnetic field. Those observations represent the birth of modern NMR spectroscopy. In 1952 Purcel and Bloch shared the Nobel Prize in physics for their discoveries [85, 86]. In the 1980s Ernst, Wüthrich and co-workers developed the methodology that enabled the determination of protein structures by developing the first two- dimensional NMR experiments, and established the Nuclear Overhauser Effect as a convenient way for measuring inter-atomic distances within proteins. The first three-dimensional structure of a small protein was solved in the lab of Kurt Wüthrich in 1985 [87]. In 1991 Richard R. Ernst was awarded with the Nobel Prize in chemistry for “for his contributions to the development of the methodology of high resolution nuclear magnetic resonance (NMR) spectroscopy” [88].

In 2002 Kurt Wüthrich received the Nobel Prize in chemistry "for his development of nuclear magnetic resonance spectroscopy for determining the three-dimensional structure of biological macromolecules in solution" [87].

Today, NMR spectroscopy represents an effective tool in physics, chemistry, biochemistry, medicine and biology. Magnetic resonance imaging (MRI) is probably the most well-known application of the NMR principle. In the life-science sector NMR has become an important method in structural biology and (structural) biophysics [89]. Beside structure determination, NMR spectroscopy is also a versatile and efficient technique for studying protein interactions as well as protein dynamics [90]. One perceived main advantage of NMR spectroscopic techniques for structural biology is the possibility to study the three dimensional structure of a biomolecule (i.e. protein) in solution and thus close to the physiological conditions that are present inside a cell. In contrast X-ray crystallographic studies, commonly used for the determination of three-dimensional protein structures, rely and depend on the availability of high quality protein crystals grown from the protein in solution by controlled precipitation. However, not all proteins do form such crystals and thus protein crystallization is still based on trial and error optimization of the crystallization process for each new target protein. At the same time, as outlined above, the crystal lattice, the dense, ordered packing of protein molecules within such crystals, might prevent or inhibit large-scale conformational changes that are of primary interest in the field of (photo)sensor structural biology.

As also outlined above, the use of NMR spectroscopic methods essentially eliminates, all the aforementioned problems of X-ray crystallography [91]. In particular no crystals need to be grown for structural studies and the protein is present in a more natural aqueous environment, although at relatively high concentrations (i.e. mM). However, one problem associated with solution NMR spectroscopy of proteins is the molecular size of the protein under investigation. With increasing size of the biomolecule, too many signals will be present and spectral resolution limits the determination of high-resolution solution protein structures. Another problem is the fact that in large proteins the magnetization relaxes faster (see below), hence there is less time to detect the signal and the corresponding peaks become broader and weaker. Although, recent developments have managed to push the size limitation considerably [92], true atomic-resolution structures of huge Mega-Dalton protein complexes such as the ribosome are still out of reach for NMR techniques.

Since NMR spectroscopy is the key technique used in the presented thesis, certain principles of NMR spectroscopy will be introduced in the following.

1.3.1 The NMR phenomenon

All matter is made up of atoms which are constituted of atomic nuclei (protons and neutrons) surrounded by a “cloud” of electrons. Whereas, X-ray crystallography employs the diffraction of X-ray radiation by the electron “cloud” of the respective atom for structure determination, NMR spectroscopy relies on certain physical properties of the atomic nucleus. Atomic nuclei have four important properties: mass, electric charge, magnetism and spin. Hereby, the latter one is of primary importance for all NMR experiments. The name spin was originally conceived to describe the rotation of a particle around some axis. For the spin of nuclei as quantum scale objects, this picture is in so far correct as spins obey the same mathematical laws as quantized angular momenta do. Classically, an angular momentum of a macroscopic particle is a vector quantity represented by the product of a particles rotational inertia and rotational velocity about a particular axis. For quantum scale objects such as nuclei this angular momentum is quantized and hence can only possess certain fixed values. The spin angular momentum is associated with a magnetic moment which can be visualized as a tiny bar magnet.

The overall spin of the nucleus is determined by the spin quantum number S . The nuclear spin (I) is important. If the number of neutrons or number of protons is odd, then the nucleus has a half integer spin (i.e. $I = \frac{1}{2}$). The spinning action of such an nucleus and it's positive charge generates a small magnetic field. Nuclei with even atomic numbers and thus $I = 0$ such as ^{12}C and ^{16}O are not detectable in NMR spectroscopy since they do not produce any magnetic field. Correspondingly, spinning nuclei with an odd atomic number possess a nuclear magnetic moment μ . To be more precise, the magnetic moment μ is directly proportional to P with a proportionality constant γ ,

$$\boxed{\mu = \gamma P} \quad (1)$$

where γ is the gyromagnetic ratio of the observed nucleus and P is angular momentum.

In a magnetic field, a nucleus of spin I will have $2I+1$ possible orientations. For nuclei with $I = \frac{1}{2}$, two magnetic quantum numbers $m = \frac{1}{2}$ and $m = -\frac{1}{2}$ are possible. Those are often denoted as α and β states of the nucleus. In the absence of an external magnetic field, these orientations are of equal energy. In the presence of an external magnetic field their energy level split (Fig. 6). According to thermodynamics the lower energy level (along the direction of magnetic field) will contain slightly more nuclei than the higher level. It is possible to excite these lower energy level nuclei into the higher energy level using electromagnetic radiation.

1.3.1.1 Energy level diagram

The energy required to excite these nuclei and to obtain an NMR signal is just the energy difference between the two nuclear orientations and depends on the strength of the magnetic field B_0 (Fig. 6).

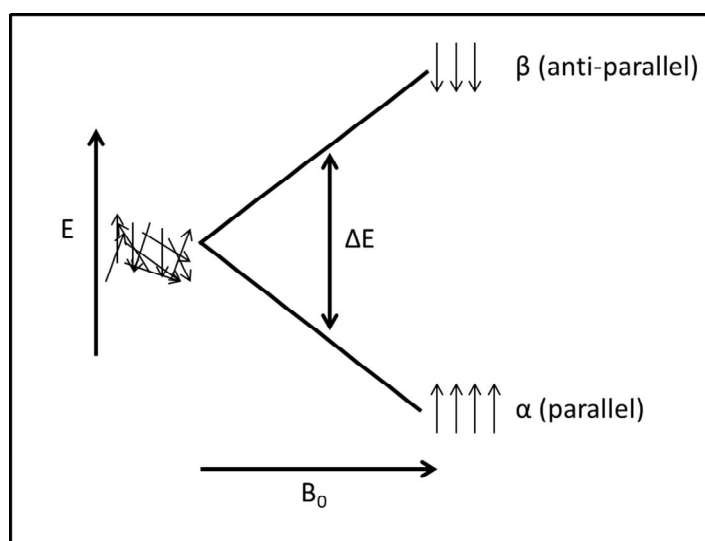


Figure 6: The energy difference (ΔE) between the α - and β -state is proportional to the external magnetic field B_0 .

1.3.1.2 Larmor frequency

The magnetic moment of each nucleus precesses around B_0 . The nuclear magnetic moments do not align with the external field, because the nuclei possess a spin angular momentum and start to rotate around the B_0 axis, much like a spinning top. The frequency of this precession is called Larmor frequency (ω),

$$\omega = \gamma B_0$$

(2)

where B_0 is the strength of the applied magnetic field and γ is the gyromagnetic ratio. The gyromagnetic ratio is the fundamental property that gives rise to a specific resonance frequency of a given isotope.

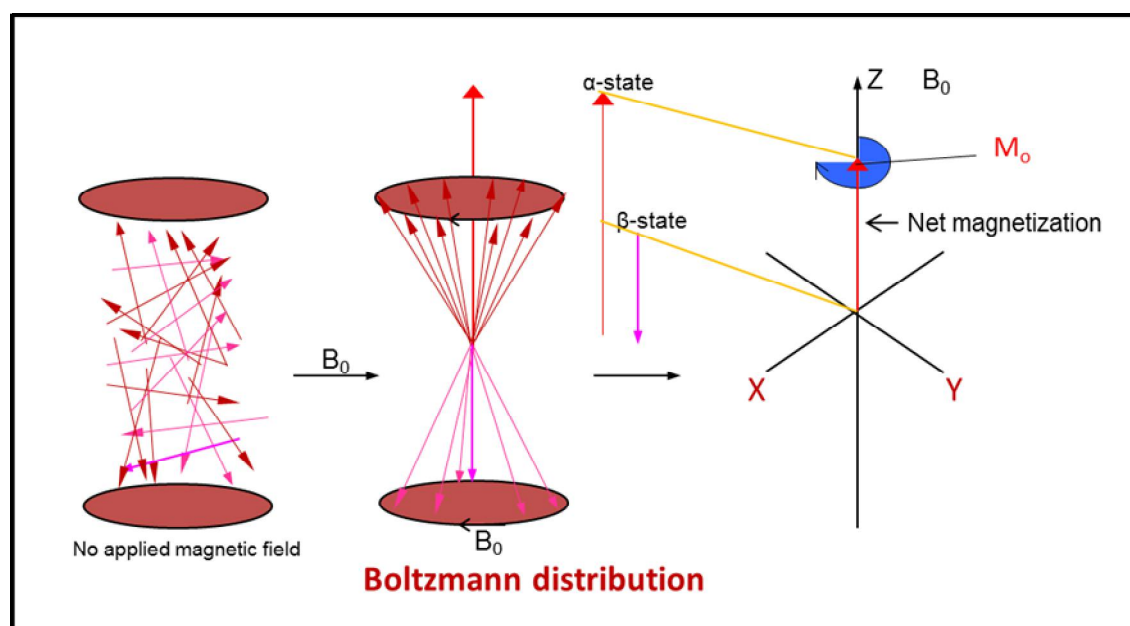


Figure 7: Alignment of the α - and β - states of nuclei in a magnetic field B_0 . The population difference between the α - and β -state is greatly exaggerated to highlight the origin of the net magnetization. No net magnetization is observed in the x, y-plane.

Distinct population of the α and β states (Fig. 6) produces a net nuclear magnetization along the external field. The net magnetization is proportional to the energy difference between α and β states.

1.3.2 Origin of the NMR signal

To record a 1D-NMR spectrum, a short radiofrequency pulse B_1 is applied and thus excites particular nuclei in a sample. At the same time the net magnetization rotates away from z-axis to the x, y plane. During back relaxation of the excited nuclear spin to its original orientation a radiofrequency signal is emitted. If multiple NMR-active nuclei are present in the sample the emitted signals are a superposition of all excited frequencies. Its evolution in time is recorded. This time-dependent signal is called free induction decay (FID). Hence, the FID contains all

the resonance frequencies of the NMR-active nuclei, corresponding to their chemical shifts. In modern NMR techniques Fourier transformation (FT) [93] of the FID yields the conventional 1D spectrum. In 1D NMR spectra, a number of peaks are present which can be described as below:

- (i) The number of signals shows how many different types of NMR active nuclei are present
- (ii) The location of the signal (chemical shift) reflects the microenvironment of the respective NMR active nucleus
- (iii) The intensity of the signal relates to the number of nuclei of the same type
- (iv) Signal splitting indicates the number of NMR active nuclei on adjacent atoms

1.3.3 Chemical shift

The magnetic field at the nucleus is not equal to the external magnetic field B_0 because the electrons around the nucleus shield it from the external magnetic field. The difference between the external magnetic field and the field at the nucleus is termed as nuclear shielding and is observed as a shift of the nucleus Larmor frequency. The frequency shift in NMR spectrum is termed chemical shift. The chemical shift is defined as nuclear shielding from the external magnetic field. The chemical shift is denoted as δ and is given in ppm (parts per million) (equation 3).

$$\delta = \frac{\omega - \omega_{ref}}{\omega_{ref}} \quad (3)$$

where ω is the resonance frequency and ω_{ref} is the reference frequency.

For example, the chemical shift reference standard for ^1H and ^{13}C NMR experiments is tetramethylsilane (TMS), whose chemical shift is referenced to 0.0 ppm.

1.3.4 Higher-dimensional NMR experiments

For the study of biomolecules such as proteins, 1D NMR spectroscopy is unable to resolve the frequency of the individual nuclear spins due to the large number of NMR active spin systems. Limitations by signal overlap can often be overcome by extending the measurement into a second dimension or third dimension. In 2D NMR experiments the information comes from the interactions between two nuclei, either through the bond (J-coupling) or through space (Nuclear Overhauser Effect, NOE) [94].

1.3.5 NMR experiments for the structural study of proteins and peptides in solution

1.3.5.1 Structure determination of small peptides in solution

For a small peptide, three types of 2D experiments are widely used for structure determination: 2D ^1H - ^1H COSY (*correlation spectroscopy*), 2D ^1H - ^1H TOCSY (*total correlation spectroscopy*) and 2D ^1H - ^1H NOESY (*nuclear overhauser spectroscopy*). In these experiments signals from the same isotope (usually ^1H which is highly abundant and nmr active) are detected. Backbone and side-chain proton resonances are assigned in a sequential manner using the 2D ^1H - ^1H COSY and 2D ^1H - ^1H TOCSY spectra. NOESY experiments are used to determine the structure of a molecule. In this experiment the coherence is transferred through the space rather than through the bond, thus the signals contain information about the distance between two NMR active nuclei and thus provide short to medium range structural restraints for structural calculations. The detection of correlation between two protons is only possible if the distance is below 6 Å. Since short peptides contain a limited number of amino acids, spectral overlap is often limited and thus unlabeled peptide samples can be used for structural studies.

1.3.5.2 Structure determination of proteins in solution

Apart from ^1H , proteins contain other NMR active nuclei e.g. ^{13}C , ^{15}N and ^{31}P . Hereby, ^{13}C and ^{15}N are of special importance. However, those NMR active carbon and nitrogen isotopes are naturally of low abundance (< 1%) [94]. Therefore, protein samples usually need to be isotopically labeled using a ^{15}N nitrogen source such as ^{15}N ammoniumchloride and/or a ^{13}C carbon source such as ^{13}C D-glucose. Labeling is performed by growing target protein

overproducing *E. coli* cells in minimal media using either one or both of the labeled compounds as sole N and C source [95]. In NMR experiments the use of those hetero nuclei allows the determination of the structure of much larger polypeptides such as proteins.

Using isotopically labelled protein samples it is possible to measure the most important heteronuclei experiment, the so-called Heteronuclear Single Quantum Coherence (HSQC) experiment. Each signal in a ^1H - ^{15}N HSQC spectrum represents a proton that is bound to a nitrogen atom. For proteins, the spectrum thus contains the backbone amide (NH) signals of the proteins polypeptide chain as well as the NH signal of certain amino acid side-chains such as glutamine and asparagine. In ^1H - ^{15}N HSQC experiments, usually all amino acids are visible except prolines, which lack a backbone amide. 2D spectra are often crowded with signals. Therefore, these spectra are extended in the third dimension, so that signals are distributed in a cube instead of a plane. Commonly such triple resonance experiments need to be measured for the sequential assignment of protein resonances. In triple resonance experiments three different nuclei (^1H , ^{15}N and ^{13}C) are correlated. Hereby, the names of the triple resonance experiments are very descriptive. The names of all nuclei, which are used for magnetization transfer during the experiment, are listed in order of their use, bracketing the names of the nuclei which are used only for transfer and whose frequencies are not detected: HNCA [96], HN(CO)CA, HNC(O) and HN(CO)CA. For details see the materials and methods section (chapter 2.11.6.1). These experiments are run in pairs with one experiment giving rise to correlations to the residue itself (residue *i*) and the previous residue (residue *i*-1) and its partner experiment giving information about the previous residue (residue *i*-1). Like for peptide samples, distance restraints are needed for 3D structural calculations. Those are provided by the corresponding ^{15}N and/or ^{13}C NOESY experiments.

1.3.5.3 NMR spectroscopy: more than a method to determine protein structures

Once the backbone assignment is complete or near to complete, the information can be used to measure NMR experiments which provide information about the dynamics of the protein. Such experiments are e.g. spin relaxation experiments and hydrogen/deuterium (H/D) exchange experiments.

Relaxation experiments:

Molecular tumbling (global) and protein motion (local) cause fluctuations in the local magnetic field that lead to relaxation. The backbone ^1H - ^{15}N heteronuclear NOE experiment provides information about the motion of individual N-H bond vectors. Those that undergo motion faster than the overall tumbling of the molecules (i.e. on the pico- to nano second timescale) show a decreased NOE intensity relative to the average observed for the majority of the residues. Thus, for instance, decreased (or negative) hetNOE values are usually found at the flexible N- and C-terminal ends of a protein. The most important relaxation processes used to determine motions in the respective experiments are:

- 1) T1 relaxation experiment: The longitudinal relaxation (T1) or spin-lattice relaxation is measured which corresponds to the relaxation of the spin from perturbed state back to the thermal equilibrium.
- 2) T2 relaxation experiment: The transverse relaxation (T2) or spin-spin relaxation, i.e. dephasing of the magnetization in the x, y- plane is measured. Loss of coherence of the precessing spins leads to loss of NMR signal.
- 3) Heteronuclear NOE experiment: The Nuclear Overhauser Effect (NOE), that is due to dipolar-dipolar cross correlation relaxation is measured, which describes how fast magnetization is transferred from one nucleus to another [97, 98]. The measured NOE is actually a cross-relaxation rate, in that the non-equilibrium states of one nucleus affects other nuclei.

H/D exchange experiments:

H/D exchange experiments in turn provide a quantitative measure for the solvent accessibility of a given amino acid in the protein. As non-hydrogen bonded protons are constantly exchanging with the solvent, they can be exchanged against deuterium (^2H or D), with the rate of exchange depending on the solvent accessibility of the respective proton. Since ^2H has a different gyromagnetic ratio than ^1H , it is invisible in ^1H - ^{15}N HSQC experiments due to the difference in its resonance frequency. This means that this setup is suitable for monitoring the amide proton exchange by recording the signal decay over time, which can vary between seconds to even days [99]. Typically in H/D exchange NMR experiments the exchange of the covalently bonded hydrogen atoms in the backbone of the protein is monitored. Ideally, the experiment provides quantitative information about the solvent accessibility for each

backbone amide of the protein molecule. Solvent accessible and hence exposed backbone amide protons will exchange faster than buried ones.

In order to perform such an experiment a ^{15}N labeled protein sample needs to be lyophilized. Immediately, before ^1H - ^{15}N HSQC spectra are recorded the sample is re-dissolved in D_2O . After dissolving the sample in D_2O , ^1H - ^{15}N HSQC spectra are recorded in a sequential manner providing information about the rate of the H/D exchange at a given backbone amide position.

1.4 Scope and outline of the thesis

As outlined in chapter 1.2.5 all so far reported X-ray crystallographic studies on LOV photoreceptor light-dependent conformational dynamics recorded on dark grown crystals revealed only minor overall structural changes after photo-activation [54, 66, 73, 75, 100, 101]. This implies that either large scale conformational changes are absent or are impaired by the crystal lattice [100]. In contrast, solution methods such as NMR- and circular dichroism (CD) spectroscopy i.e. on the oat phot1-LOV2 and other plant LOV domains suggested pronounced light-dependent structural changes in the sensor domain as well as in auxiliary structural regions which flank the sensor core and provide in the full-length photoreceptor a connection to eventually fused effector domains (see chapter 1.2.6). These observations stress the importance of further structural studies on full-length LOV photoreceptor proteins by NMR spectroscopy and other solution techniques. Any obtained structural information will be valuable for both the understanding of LOV domain signaling and at the same time could provide a rational basis for the engineering of LOV-domain based optogenetic switches for biotechnological/biomedical purposes.

The presented thesis thus focuses on the two short LOV proteins PpSB1-LOV and PpSB2-LOV from *P. putida* KT2440, which as simplistic full-length prototypic photoreceptor proteins yet lack a fused effector domain. As sequence and architectural similarities (i.e. N-terminal and C-terminal extensions) exist to both the AsLOV2 and the YtvA photoreceptor system, the *P. putida* LOV proteins represent a particularly interesting target to study the conservation of intra-molecular signal-relay mechanisms between the pro- and eukaryotic domains of life.

Methodologically, solution NMR spectroscopy as well as complimentary biochemical and biophysical methods will be used as a tool to:

- i) study the **role and function of the PpSB1-LOV and PpSB2-LOV C-terminal Ja-extensions** and ...
- ii) characterize the light-induced **conformational changes associated with PpSB1-LOV photo-activation**.

Chapter 2: Materials and Methods

2.1.1 Bacterial strains

All bacterial strains used in this study are listed in Table 3.

Table 3: Genotype of the *E. coli* strains used in this study

Strain	Genotype	Reference
<i>E. coli</i> BL21 (DE3)	<i>F⁻ ompT hsdSB (rB⁻ mB⁻) gal dcm (λIts857indI Sam7 nin5 P_{lacUV5}-T7gene1)</i>	Studier & Moffatt, 1986 [102]
<i>E. coli</i> DH5α	<i>supE44Δ(lacZYA argF)U196(Φ80ΔlacZM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	Woodcock <i>et al.</i> , 1989 [103]
<i>E. coli</i> CmpX131 (C41(DE3) Derivative)	<i>F⁻ ompT hsdSB (rB⁻ mB⁻) gal dcm (λIts857indI Sam7 nin5 P_{lacUV5}- T7gene1) ribCA ribM</i>	Mathes <i>et al.</i> , 2009 [104]

2.1.2 Plasmids

All plasmid vectors used for cloning and heterologous expression are listed in Table 4.

Table 4: Plasmids used in this study

Plasmid	Genotype	Reference
pET28a	<i>ColE1 lacZ' Kan^R P_{T7} P_{lac}</i>	Novagen
pET28a-PpSB1-LOV	426 bp DNA fragment encoding for PpSB1-LOV (UniProt-Q88E39) in pET28a; <i>Kan^R</i>	Jentzsch <i>et al.</i> , 2009 [49]
pET28a-PpSB1-LOV-ΔJα	357 bp DNA fragment encoding for PpSB1-LOV without C-terminal Jα-helix (residue 1 – 119) in pET28a; <i>Kan^R</i>	This study
pET28a-PpSB1-LOV-Δcap	384 bp DNA fragment encoding for PpSB1-LOV without N-terminal cap (residue 15 – 142) in	This study

pET28a; <i>Kan^R</i>		
pET28a-PpSB1-LOV-core	315 bp DNA fragment encoding for PpSB1-LOV without C-terminal J α -helix and N-terminal cap (residues 15 – 119) in pET28a; <i>Kan^R</i>	This study
pET28a-PpSB2-LOV	444 bp DNA fragment encoding for PpSB2-LOV (UniProt-Q88JB0) in pET28a; <i>Kan^R</i>	Jentsch <i>et. al</i> 2009 [49]
pET28a-PpSB2-LOV-ΔJα	357 bp DNA fragment encoding for PpSB2-LOV without C-terminal J α -helix (residue 1 – 119) in pET28a; <i>Kan^R</i>	This study
pET28a-PpSB2-LOV-Δcap	402 bp DNA fragment encoding for PpSB1-LOV without N-terminal cap (residue 15 – 148) in pET28a; <i>Kan^R</i>	This study
pET28a-PpSB2-LOV-core	315 bp DNA fragment encoding for PpSB1-LOV without C-terminal J α -helix and N-terminal cap (residues 15 – 119) in pET28a; <i>Kan^R</i>	This study

2.1.3 Oligonucleotide primers

Sequences of oligonucleotide primers used for PCR amplification and sequencing reactions are listed in Table 5.

Table 5: Oligonucleotide primers used in this study

Name	Sequence (5' - 3') (restriction endonuclease cleavage site <u>underlined</u>) [§]	T _m (°C)
PpSB1-LOV-JαFor	5'-TCCAGTAC <u>CATATG</u> ATCAACGCGCAATTGCTGC-3'	63.6
PpSB1-LOV-JαRev	5'-TACAAT <u>CTCGAGT</u> CAGACGTCCTTCTGGATGC-3'	63.3
PpSB1-LOV-ΔcapFor	5'-TCCAGTAC <u>CATATG</u> GACGGCATCGTGGTTGC-3'	64.7
PpSB1-LOV-ΔcapRev	5'-TACAAT <u>CTCGAGT</u> CAGGCGCGTTCGTCGG-3'	66.2
PpSB1-LOV-coreFor	5'-TCCAGTAC <u>CATATG</u> GACGGCATCGTGGTTGC-3'	64.7
PpSB1-LOV-coreRev	5'-TACAAT <u>CTCGAGT</u> CAGACGTCCTTCTGGATGC-3'	63.3
PpSB2-LOV-JαFor	5'-CCTACAAT <u>CATATG</u> ATCAACGCAAACTCCTGC-3'	60.7
PpSB2-LOV-JαRev	5'-CGCTAA <u>CTCGAGT</u> CAGACATCGCGCTGGATGC-3'	67.2
PpSB2-LOV-ΔcapFor	5'-ATTATAAT <u>CATATG</u> GATGGCATCGTTGTCCGAGC-3'	63.6
PpSB2-LOV-ΔcapRev	5'-CGCTAA <u>CTCGAGT</u> CAGTGCTTGGCCTGG-3'	65.8
PpSB2-LOV-coreFor	5'-ATTATAAT <u>CATATG</u> GATGGCATCGTTGTCCGAGC-3'	63.6
PpSB2-LOV-coreRev	5'-CGCTAA <u>CTCGAGT</u> CAGACATCGCGCTGGATGC-3'	67.2

§: All forward (For) oligonucleotide primers contained a *Nde*I and all reverse (Rev) oligonucleotides contained a *Xho*I restriction endonuclease cleavage site.

2.2 Chemicals

All chemicals used in this study are listed in Table 6 and were obtained in the highest purity available.

Table 6: Chemicals used in this study

Chemical	Supplier
2,2,2-Trifluoroethanol (TFE)	Sigma Aldrich (Steinheim, Germany)
[U-¹³C₆, 99 %] D-Glucose	Euriso-Top (Saint-Aubin Cedex, France)
[U-¹⁵N, 99 %] Ammoniumchloride	Euriso-Top (Saint-Aubin Cedex, France)
Acrylamid / bisacrylamide mixture, Rotiphorese® Gel 30 (37.5:1)	Roth (Karlsruhe, Germany)
Adenosintriphosphat, ATP	Fermentas (St. Leon-Rot, Germany)
Agar-Agar	Roth (Karlsruhe, Germany)
Ampicillin	Gerbu (Heidelberg, Germany)
Ammoniumpersulfate	Pharmacia (Munich, Germany)
β-Mercaptoethanol	Sigma Aldrich (Steinheim, Germany)
Bromophenol blue	Serva (Duisburg, Germany)
Calciumchlorid	Roth (Karlsruhe, Germany)
Cobalt(II) chloride (CoCl₂)	Roth (Karlsruhe, Germany)
Copper(II) chloride (CuCl₂)	Roth (Karlsruhe, Germany)
Deuterium oxide (99,99 %)	Sigma Aldrich (Steinheim, Germany)
Dithiothreitol (DTT)	Sigma Aldrich (Steinheim, Germany)
Ethylenediaminetetraacetic acid (EDTA)	Merck (Darmstadt, Germany)
Ferric chloride(FeCl₃)	Roth (Karlsruhe, Germany)
Flavin mononucleotide (FMN)	Sigma Aldrich (Steinheim, Germany)
Glacial acetic acid	Roth (Karlsruhe, Germany)
Glycerol	Roth (Karlsruhe, Germany)

Imidazole	Roth (Karlsruhe, Germany)
Isopropanol	Roth (Karlsruhe, Germany)
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	Roth (Karlsruhe, Germany)
Kanamycin	Gerbu (Heidelberg, Germany)
Manganese(II) chloride(MnCl₂)	Roth (Karlsruhe, Germany)
Monosodium phosphate (NaH₂PO₄)	Applichem (Darmstadt, Germany)
Monopotassium phosphate (KH₂PO₄)	Roth (Karlsruhe, Germany)
Protease inhibitor (Complete, EDTA free)	Roche (Freiburg, Germany)
Protein molecular weight marker	Fermentas (St. Leon-Rot, Germany)
Sodium chloride	Roth (Karlsruhe, Germany)
Sodium dodecyl sulfate (SDS)	Serva (Duisburg, Germany)
Nickel(II) chloride (NiCl₂)	Roth (Karlsruhe, Germany)
Ni²⁺ - Nitrilotriacetic acid (Ni²⁺-NTA)	Qiagen (Hilden, Germany)
N,N,N',N'-Tetramethylethyldiamin (TEMED)	Pharmacia (Munich, Germany)
Riboflavin	Sigma Aldrich (Steinheim, Germany)
Sodium dodecylsulfate (SDS)	Serva (Duisburg, Germany)
Sodium molybdate (Na₂MoO₄)	Roth (Karlsruhe, Germany)
Streptavidin resin	Merck Millipore (Darmstadt, Germany)
Tris (hydroxymethyl) aminomethane (Tris)	Pharmacia (Munich, Germany)
Tryptone	Roth (Karlsruhe, Germany)
Yeast extract	Roth (Karlsruhe, Germany)
Zinc sulfate (ZnSO₄)	Roth (Karlsruhe, Germany)

2.3 Laboratory equipment

All laboratory equipment as well as analytical instruments used in this study are listed in Table 7.

Table 7: Laboratory instruments

Instrument	Product	Supplier
ÄKTApurifier	Purifier 100 Plus	GE Healthcare (Freiburg, Germany)
CD spectrometer	JASCO J-810	Jasco (Groß-Umstadt, Germany)
Centrifuge	Biofuge pico	Heraeus (Hanau, Germany)
Centrifuge	Rotina 35R	Hettich (Kuppenheim, Germany)
Centrifuge	RC-5B/RC-5B plus	Dupont Sorvall (Apeldoorn, Netherlands)
French Press	French Press 40K Cell	Thermo Electron Corporation (Essen, Germany)
Freeze-dryer	Lyovac GT2	SRK System Technik (Riedstadt, germany)
HPLC	RF-LC-2010 Ai	Shimadzu (Duisburg, Germany)
HPLC Column	250/4 Nucleodur C18ec	Macherey&Nagel (Düren)
MALDI-TOF MS	Voyager DE Pro MALDI-TOF	Applied Biosystems (Foster City, CA, USA)
NMR spectrometer	600MHz	Varian (Lake Forest, California, USA)
NMR spectrometer	800MHz	Varian (Lake Forest, California, USA)
NMR spectrometer	900MHz	Varian (Lake Forest, California, USA)
PCR Thermocycler	Infors HT	Infors (Einsbach, Germany)
pH-Meter	766 Calimetric	Knick (Berlin, Germany)
SDS-PAGE apparatus	Mini-PROTEAN Gel-chamber II	Bio-Rad (Munich, Germany)

SDS-PAGE apparatus	XCell SureLock Mini-Cell	Sartorius (Göttingen, Germany)
SDS-PAGE Documentation	Stella	Raytest (Straubenhardt, Germany)
Spectrophotometer	DU 650	Beckmann (Krefeld, Germany)
Thermomixer	Comfort	Eppendorf (Hamburg, Germany)

2.4 Software and Databases

Protein sequence database enquiries were carried out using the European Molecular Biology Laboratory (EMBL)- and GenBank database (www.ebi.ac.uk/embl/Access/index, www.ncbi.nlm.nih.gov/genbank). Protein structure database enquiries and depositions were carried out using Protein Data Bank Europe (www.ebi.ac.uk/pdbe). Nuclear magnetic resonance assignments have been deposited at the Biological Magnetic Resonance Bank (www.bmrb.wisc.edu).

Table 8: Software used in this study

Software product	Reference
ARIA2	Rieping <i>et al.</i> , 2007 [105]
CcpNmr Analysis	Vranken <i>et al.</i> , 2007 [106]
ClustalW2	Thompson <i>et al.</i> , [107]
Common Interface for NMR structure Generation (CING)	Doreleijers <i>et al.</i> , 2012 [108]
Crystallography and NMR System (CNS) 1.2.1	Brunger <i>et al.</i> , 2007 [109]
ExPASy Molecular Biology server	Gasteiger <i>et al.</i> , 2003 [110]
Molprobit 3.1.8	Davis <i>et al.</i> 2007 [111]

NMRPipe 5.5.X	Delaglio <i>et al.</i> 1995 [112]
ORIGIN 7	Microcal Software Inc. (Northhampton, USA)
Pymol	DeLano <i>et al.</i> , 2002 [113]
TALOS+	Shen <i>et al.</i> 2009 [114]

2.5 Microbial media and antibiotics

All media components were autoclaved for 20 min at 120 °C at a constant pressure of 200 kPa. Heat-labile components such as antibiotics, vitamins, riboflavin, and trace metal ions were either dissolved in autoclaved dH₂O or were filter sterilized (Schleicher & Schuell, pore size 0.2 µm).

2.5.1 Antibiotics

Kanamycin (50 mg/ml stock solution in dH₂O, 50 µg/ml working concentration)

2.5.2 Full media for production of large amounts of recombinant proteins

2.5.2.1 Luria-Bertani (LB) medium [115]

10 g/l bacto-tryptone

5 g/l yeast extract

10 g/l NaCl

Adjust volume with dH₂O

2.5.2.2 Autoinduction full media (AI-FM) (modified according to Studier *et al.*) [116]

Media component:

12 g casein hydrolysate

24 g yeast extract

5 g glycerol

Adjust the volume to 800ml with MilliQ dH₂O.

Buffer Component:

53.3 g K₂HPO₄

26.4 g KH₂PO₄

Mix together and adjust the volume to 500 ml with MilliQ dH₂O.

Induction Component:

5 % (w/v) glucose

2 % (w/v) lactose

All media components were autoclaved separately.

To obtain 1 liter of autoinduction media, 800 ml of the media component, 90 ml of buffer component, 10 ml of glucose solution and 100 ml of lactose solution were mixed.

Please note: For the growth of the riboflavin auxotroph *E. coli* CmpX131 (DE3) (2.9.1.2) strain all media had to be supplemented with 50 μM of riboflavin [104].

2.5.3 Minimal media used for the production of isotopically labeled proteins

2.5.3.1 HD-MM (High cell density IPTG induction) minimal medium [117]

50 mM Na₂HPO₄·7H₂O

25 mM KH₂PO₄

10 mM NaCl

5 mM MgSO₄

0.2 mM CaCl₂,

0.25x trace element mixture

0.25x vitamin stock solution

0.1 % (w/v) NH₄Cl or [U-¹⁵N, 99%] NH₄Cl,

1.0 % (w/v) D-Glucose or [U-¹³C₆, 99 %] D-Glucose

Adjust volume with MilliQ dH₂O

Trace element mixture (1000x) [116]:

100 mM FeCl₃ (dissolved in 0.12 M HCL)

50 mM FeCl₃, 20 mM CaCl₂

10 mM MnCl₂

10 mM ZnSO₄

2 mM CoCl₂

2 mM CuCl₂

2 mM NiCl₂

2 mM Na₂MoO₄.

Adjust volume with MilliQ dH₂O

Vitamin stock solution [116]:

20 mg/l Thiamine

20 mg/l Biotin

Adjust volume with MilliQ dH₂O

2.5.3.2 Autoinduction minimal media (AI-MM medium) [117]

50 mM Na₂HPO₄

50 mM KH₂PHO₂ (pH 8.2)

50 mM NH₄Cl

5 mM Na₂SO₄

2 mM MgSO₄

Trace element mixture (1000x) (2.5.3.1)

0.75 % (v/v) glycerol

0.05 % glucose

0.01 % lactose

10 mg/l Thiamine

Adjust volume with MilliQ dH₂O

Please note: For the growth of the riboflavin auxotroph *E. coli* CmpX131 (DE3) (2.9.1.2) strain all media had to be supplemented with 50 µM of riboflavin [104].

2.5.3.2.1 Preparation of agar plates

For the preparation of agar-plates the respective liquid medium was supplemented with 1.5 % (w/v) agar-agar.

2.6 Buffer and solutions

2.6.1 Agarose gel-electrophoresis

Running buffer:

10x TBE buffer

89 mM Tris

89 mM Boric acid

2.5 mM EDTA

6x DNA loading dye (Working concentration: 1x):

0.09 % (w/v) bromphenol blue

0.09 % (w/v) xylene cyanol FF

60 % (v/v) glycerol

60 mM EDTA.

Ethidium bromide stock solution:

10 mg/ml stock solution;

Working concentration: 10⁵ dilution of stock solution

DNA molecular weight standard:

“GeneRuler 100 bp DNA Ladder“ (Fermentas) 1000; 900; 800; 700; 600; 500; 400; 300; 200; 100 bp. (GeneRuler 1 kb DNA Ladder, Fermentas).

2.6.2 CaCl₂ solution for preparation of chemically competent *E. coli* cells

60 mM CaCl₂

15 % (w/v) glycerol

10 mM PIPES

pH adjusted to 7.0

2.6.3. Buffers for protein purification

2.6.3.1 Lysis Buffer

10 mM sodium phosphate buffer (pH 8.0),

10 mM NaCl

1 mM DTT

Complete Protease inhibitor (EDTA free) (1 tablet/L) (Roche, Freiburg, Germany)

2.6.3.2 Ni-NTA Chromatography buffers

Ni-NTA equilibration buffer:

10 mM sodium phosphate buffer (pH 8.0),

10 mM NaCl

1 mM DTT

Ni-NTA wash buffer:

10 mM sodium phosphate buffer (pH 8.0)

10 mM NaCl,

1 mM DTT

20 mM imidazole

Ni-NTA elution buffer:

10 mM sodium phosphate buffer (pH 8.0)

10 mM NaCl

1 mM DTT

250 mM imidazole.

2.6.3.3 Dialysis, NMR and protein storage buffer

10 mM sodium phosphate buffer (pH 6.5)

10 mM NaCl

1 mM DTT

2.6.4 Sodium-dodecyl polyacrylamide gel-electrophoresis (SDS-PAGE)

[118]

Stacking gel buffer:

0.5 M Tris/HCl pH 6.5

Separation gel buffer:

1.5 M Tris/HCl pH 8.5

SDS running buffer (10x):

3 % (w/v) Tris base

1.29 M Glycine

1 % SDS

Protein loading buffer 4x:

200 mM Tris/HCl, pH 8.0

8 % (w/v) SDS

8 % (v/v) β -mercaptoethanol

10 % (v/v) Glycerol

0.05% Bromophenol blue

Colloidal coomassie gel staining solution [119]:

85 % (v/v) *o*-phosphoric acid,

750 mM Ammonium sulphate

1.4 mM Coomassie G250

20 % (v/v) Methanol.

Adjust volume with dH₂O

Protein molecular weight standard:

“PageRuler™ Unstained Protein Ladder“, Fermentas: 10; 15; 20; 25; 30; 40; 50; 60; 70; 85; 100; 120; 150; 200 kDa

2.6.5 Proteolytic removal of hexa-histidine-tag

Thrombin digestion buffer:

20 mM Tris-HCl, pH 8.4

150 mM NaCl

2.5 mM CaCl₂.

2.6.6 Buffer for HPLC-SEC experiments

HPLC-SEC running buffer for proteins:

100 mM sodium phosphate buffer (pH 6.5)

10 mM NaCl

1 mM DTT

HPLC-SEC running buffer for peptides:

200 mM sodium phosphate (pH 7.5)

10 mM NaCl

2.6.7 Bradford Reagent

0.1 g Coomassie Brilliant Blue G-250

100 ml 85 % (v/v) *o*-phosphoric acid

50 ml Ethanol

Adj. 1 liter dH₂O

The mixture is stirred overnight, filtered and stored in a dark bottle at 4 °C.

2.7 Enzymes

DNA modifying enzymes such as T4-Ligase, Taq- and Pfu Polymerase as well as standard restriction endonucleases were obtained from Fermentas (St. Leon-Rot, Germany). Biotinylated Thrombin was purchased from Novagen (Darmstadt, Germany).

2.8 Peptide synthesis

The PpSB1-J α (residues: 120-142 of the full-length protein) and PpSB2-J α peptides (residues: 120-148 of the full-length protein) were purchased in N-acetylated custom synthesized form (98% purity) from Peptide 2.0 Inc. (Chantilly, VA, USA).

2.9 Methods

2.9.1 Microbiological Methods

2.9.1.1 Cryo-preservation and long-term storage of *E. coli* strains

100 μ l (25 % v/v) of glycerol and 100 μ l of an overnight grown culture of the respective strain were mixed and stored at -80 °C.

2.9.1.2 Growth and storage of bacterial strains

For short term storage (up to 1 week at 4 °C) bacteria were grown and stored on LB agar plates (2.5.3.2.1). *E. coli* strains were grown overnight at 37 °C. For plasmid maintenance plates were supplemented with the appropriate antibiotics (2.5.1).

Long-term storage of *E. coli* strains was carried out by cryo-preservation (2.9.1.1). *E. coli* BL21 (DE3) and *E. coli* DH5 α strains were grown overnight at 37 °C in a rotary shaker (120 rpm) in LB medium (2.5.2.1) supplemented with the appropriate antibiotics for plasmid maintenance. *E. coli* CmpX131 (DE3) strains were grown overnight in LB medium (2.5.2.1) supplemented with the appropriate antibiotics and 50 μ M riboflavin at 37 °C in a rotary shaker (120 rpm).

2.9.1.3 Preparation of competent cells by CaCl₂ method [120]

2 ml of an overnight seed-culture of the appropriate *E. coli* strain was used to inoculate 200 ml LB medium (2.5.2.1) in a sterile flask. The culture was grown at 37 °C on a rotary shaker (250 rpm) to an OD₆₀₀ of 0.4-0.6. Subsequently cultures were kept on ice for about 30 min. The cells were centrifuged for 10 min at 3000 rpm, the supernatant was removed and the cell pellet resuspended in 20 ml of ice-cold 0.1 M CaCl₂ solution (2.6.2). The resuspended cells were kept on ice for another 30 min. Finally, cells were pelleted at 4 °C for 10 min at 3000 rpm. Cells were resuspended in 2 ml of ice-cold 0.1 M CaCl₂ solution (2.6.2). Aliquots of 100 µl were stored at -80 °C until further use.

Please note: For the preparation of competent cells of the riboflavin auxotroph *E. coli* CmpX131 (DE3) strain all media had to be supplemented with 50 µM of riboflavin [104].

2.9.1.4 Heat-shock transformation of chemically competent cells [120]

E. coli cells were rendered competent by the CaCl₂ method (2.9.1.3) [120]. 50-100 ng of DNA was added to 100 µl of competent cells previously thawed on ice. Cells and DNA were incubated on ice for 30 min and subsequently heat shocked at 42 °C for 90 s. Subsequently, 1 ml of LB medium without antibiotics was added. Cells were incubated at 37 °C for 45 min under constant agitation (120 rpm). Finally, the cells were pelleted by centrifugation (14000 rpm, 3 min), the pellet resuspended in 50 – 70 µl LB medium (2.5.2.1), and subsequently plated on LB agar plates (2.5.3.2.1) containing the appropriate antibiotics for selection.

Please note: For the transformation of competent cells of the riboflavin auxotroph *E. coli* CmpX131 (DE3) strain all media had to be supplemented with 50 µM of riboflavin [104].

2.9.1.5 Induced gene expression in *E. coli*

All recombinant genes were expressed in either *E. coli* BL21 (DE3) or *E. coli* CmpX131 (DE3). A single colony of the respective *E. coli* strain harboring the respective plasmid was used to inoculate a 5 ml LB seed culture containing the appropriate antibiotics. The culture was grown overnight at 37 °C under constant agitation (120 rpm). From this seed culture a bigger culture volume of an appropriate full- or minimal media (see 2.9.1.5.1 – 2.9.1.5.4) was inoculated and grown as described below.

2.9.1.5.1 Expression of PpSB1-LOV/PpSB2-LOV wild-type and truncated constructs

1 ml of an overnight seed culture (2.9.1.5) was used to inoculate 100 ml of AI-FM (2.5.2.2). All other steps were identical to the protocol described in the following except that the temperature after induction was set to 30 °C and the culture was induced overnight.

2.9.1.5.2 Expression optimization for the isotopic labeling of proteins for NMR experiments

For the evaluation of protein production in different minimal media 1 ml of an overnight seed culture (2.9.1.5) was used to inoculate 100 ml of HD-MM medium (2.5.3.1) or AI-MM medium (2.5.3.2) (containing the conventional N- and C-sources). All other steps were identical to the protocol described in the following which was used for the production of isotopically labeled protein samples.

2.9.1.5.3 Protein production in HD-MM (2.5.3.1) medium

For the production of isotopically labeled protein samples, the overnight seed-culture was centrifuged at 5000 rpm and the pellet was resuspended in the respective minimal medium containing the isotopically labeled N- and C-sources instead of the conventional ones (2.5.3.1). For induction of gene expression in HD-MM medium, cell growth was monitored by measuring the optical density of the culture at 600 nm (OD_{600nm}) using a Beckmann DU650 UV/Vis spectrophotometer (Beckmann Coulter, Krefeld, Germany). At an OD_{600nm} of about 0.6-0.8 (for CmpX131 (DE3) strain OD_{600nm} 1.2-1.5) gene expression was induced by addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After induction cells were grown for up to 48 hours at 20 °C in the dark at constant agitation (120 rpm).

2.9.1.5.4 Protein production in AI-MM (2.5.3.2) medium

In AI-MM medium induction of gene expression occurs automatically, when cell growth exceeds a certain limit, i.e. when all of the glucose contained in the medium has been utilized. As long as glucose is still present in the medium the *Plac_{UV5}* promoter dependent expression of the T7-RNA polymerase of the employed strains is inhibited by catabolite repression [116]. After complete consumption of glucose, the strain will start to utilize the lactose present in the medium which at the same time induces the expression of the T7 polymerase and hence also of the target gene localized on a pET28a vector under control of the strong P_{T7} promoter.

During this growth phase the strain also utilizes glycerol as C-source which is present in excess in the medium. To achieve isotopic labeling of the overproduced protein ^{13}C -glycerol has to be supplied along with ^{13}C -glucose [117]. For heterologous expression, cells are grown in the dark for about 3 hours at 37 °C, shifted to 20 °C and grown for up to another 72 hours under constant agitation (120 rpm)

Subsequently the cells were harvested by centrifugation at 5000 rpm at 4 °C for 15 min. The cell pellets were stored at -20 °C until further use.

2.9.2 DNA methods

2.9.2.1 Isolation of Plasmid DNA

2.9.2.1.1 Small and large-scale plasmid preparation [115]

Plasmid DNA was prepared according to standard laboratory methods using commercially available kits.

Small scale plasmid preparations were carried out using 5 ml of overnight grown *E. coli* DH5 α cultures. Plasmid DNA was prepared using the innuPrep Plasmid Mini Kit (Analytik Jena GmbH, Jena, Germany) according to the instructions provided in the manual of the manufacturer. Plasmid DNA was eluted using pre-heated (70 °C) dH₂O.

For larger scale plasmid preparations 100 ml of an overnight grown *E. coli* DH5 α culture were used. Plasmid DNA was prepared using the illustraTM Plasmid Midi Flow kit (GE Healthcare, Munich, Germany) according to the instructions provided in the manual of the manufacturer. Plasmid DNA was eluted using pre-heated (70 °C) dH₂O.

2.9.2.2 Purification of DNA from agarose gels

Hydrolyzed plasmid DNA and/or PCR fragments were purified from agarose gel slabs using the 5Prime Agarose GelExtract Mini Kit (5PRIME, Hamburg Germany), following the instructions given in the manual of the manufacturer. Hydrolyzed plasmid DNA was purified from 1 % (w/v) agarose gels and PCR fragments from 1.5 - 2.0 % (w/v) agarose gels. DNA was eluted using pre-heated (70 °C) dH₂O.

2.9.2.3 Agarose gel-electrophoresis [115]

Agarose gel-electrophoresis is a method to separate DNA and RNA fragments according to their size (molecular weight). DNA is loaded onto an agarose gel which is placed in a vertical tank filled with running buffer (2.6.1). Subsequently an electric field is applied which results in the movement of the negatively charged DNA molecules through the pores of the agarose gels. Fragments of low molecular weight migrate faster than heavier ones. Agarose gels were prepared by melting 1 % - 2 % (w/v) agarose in 1x TBE buffer (2.6.1). For staining of DNA fragments 25 µl/L ethidium bromide was added directly into the gel. The DNA samples are mixed with 1/6 volume of 6x DNA loading dye (2.6.1) and loaded onto the gel. The rate of migration is affected by a number of factors e.g concentration of agarose, voltage applied and the conformation of the DNA. Typically gels were run at constant voltage of 100 V, removed from the tank and visualized/photographed under UV light using a EagleEye II® (Stratagene La Jolla, CA, USA) UV transilluminator.

2.9.2.4 Recombinant DNA techniques

2.9.2.4.1 Restriction of DNA by restriction endonucleases [121]

1 - 2 µg of DNA was digested using 1 - 2 units of appropriate restriction endonucleases by employing the buffer and temperature conditions suggested by the enzyme manufacturer (Fermentas, St.Leon-Rot, Germany). Digested DNA fragments (plasmid DNA and PCR products) were separated by agarose gel-electrophoresis (2.9.2.3). The desired bands were excised from the gel using a clean scalpel. DNA was purified from the agarose gel pieces as described above (2.9.2.2).

2.9.2.4.2 DNA ligation [122]

Hydrolyzed plasmid DNA and the respective PCR products were joined in a T4 ligase (Fermentas, St. Leon-Rot, Germany) dependent reaction. T4 ligase catalyzes the formation of a phosphodiester bond between the 3' hydroxyl end of one DNA molecule and the 5' phosphate of another. DNA ligations were carried out in 10 - 50 µl reaction volumes containing the respective plasmid and insert DNA in a 1:5 molar ratio. Reactions were supplemented with T4 ligase buffer (1x) and 1-5 units of T4 ligase and incubated overnight at 4 °C. Prior transformation (2.9.1.4) into an appropriate *E. coli* host strain, T4 ligase was inactivated by incubation at 65 °C for 10 minutes.

2.9.2.5 Polymerase chain reaction (PCR) [123]

DNA fragments were amplified from plasmid DNA using standard PCR techniques [123]. Typically PCR reactions (25 or 50 μ l) contained the following:

Template DNA (20-30ng)	20-30 ng
Forward primer (25 pmol/ μ l)	50 pmol
Reverse primer (25 pmol/ μ l)	50 pmol
10x polymerase buffer	1x
dNTP mix(10 mM)	0.2 mM of each dNTP
<i>Pfu</i> polymerase	1 unit

The following temperature program was employed for PCR amplification using a TProfessional Basic Gradient Thermocycler (Biometra, Göttingen, Germany).

1. Initial denaturation:	5 min,	95 °C
2. cycle denaturation:	30 s,,	95 °C
3. primer annealing	45 s,	55 °C
4. elongation	1 min	72 °C
5. cyclic repetition of steps 2-4, 30-times		
6. final elongation	10 min	72 °C

2.9.2.6 DNA Sequencing

Sequencing of plasmid DNA was performed according to the Sanger method [124] by either SeqLab GmbH (Göttingen, Germany) or Sequiserve GmbH (Vaterstetten, Germany). All pET28a-based constructs generated in this study were sequenced using standard T7 and T7-Terminator specific sequencing oligonucleotides.

2.9.2.7 Determination of DNA concentration purity

The concentration of DNA samples was estimated from agarose gels by comparison to a molecular weight standard containing defined amounts of DNA (2.6.1).

2.9.3 Biochemical and protein analytical methods

2.9.3.1 Purification of heterologously expressed proteins

In this study all proteins (unlabeled and isotopically labeled ones) were expressed as N-terminally hexa-histidine-tagged fusion proteins in *E. coli* as described in 2.9.1.5.3. This allows the easy purification of the respective constructs by immobilized metal ion affinity chromatography (IMAC) using Ni-NTA resin (see below for details).

2.9.3.1.1 Cell lysis

For cell lysis, up to 10 g of cells were thawed on ice and suspended in about 50 ml lysis buffer (2.6.3.1) containing 1 tablet of CompleteTM protease inhibitor (Roche, Freiburg, Germany). Hereby the cell suspension should not contain more than 20 % (w/v) of cells. Cells are broken by passing the cell suspension (25 ml at a time) 4 times through a chilled french pressure cell (Thermo Electron Corporation, 40K Cell) at a constant pressure of 1500 psi (300 bar cells internal pressure). Cell debris and unbroken cells were removed by centrifugation at 4 °C for 45 min at 14000 rpm. The resulting, soluble target-protein-containing supernatant (cell free lysate) was subsequently loaded onto a Ni-NTA column and purified as described below.

2.9.3.1.2 IMAC chromatography on Ni-NTA resin

The cell free lysate contains all the components which are present in the cell. Hence it is necessary to separate the target protein from the remaining cellular constituents. Hereby, the fusion with an N-terminal hexa-histidine-tag facilitates this process. Histidine, i.e. the deprotonated imidazolium ring of the histidine side-chain, possesses high affinity for Ni²⁺ ions because of the interaction between the divalent metal ion and the ring system. Recombinant proteins fused with a tag containing multiple histidines thus show affinity towards materials such as nitrilotriacetic acid complexed with Ni²⁺ ions (Ni-NTA). The use of polymeric (agarose) chromatographic materials with covalently attached Ni-NTA molecules thus facilitates efficient binding of the histidine tagged target protein and thus separation of the target protein from the remaining cellular constituents. For elution, commonly a molecule

which has higher affinity for Ni²⁺ than histidine, such as imidazole, is used to displace the target protein from the column. All buffers employed for Ni-NTA purification are listed in 2.6.3.2. Medium-scale purification of isotopically labeled proteins (as well as of their unlabeled counterparts) was carried out by gravity flow. 10 ml of a 50 % (w/v) Ni-NTA Superflow™ slurry (QIAGEN, Hilden, Germany) were placed into a gravity flow column resulting in a bed volume of 5 ml. The column was equilibrated using 10 column volumes of Ni-NTA equilibration buffer (2.6.3.2). Subsequently, the cell free lysate obtained from *E. coli* expression cultures by cell lysis (2.9.3.1.1), was loaded onto the column. The column was washed with 10 column volumes of Ni-NTA wash buffer (2.6.3.2) and bound target protein was eluted with 5 column volumes of Ni-NTA elution buffer (2.6.3.2).

2.9.3.2 Buffer exchange by dialysis

For buffer exchange of protein samples 10 kDa cutoff dialysis tubes (Float-A-Lyzer, SpectrumLabs, Breda, Netherlands) were used. Peptide samples were dialyzed using 1 kDa cutoff dialysis tubes (GE Healthcare, Munich, Germany). Samples were dialyzed overnight at 4 °C against 400 - 600 volumes of the desired buffer.

2.9.3.3 Proteolytic removal of the hexa-histidine-tag by Thrombin

The serine-protease Thrombin recognizes the amino acid sequence Leu-Val-Pro-Arg↓Gly-Ser (cleavage site shown as arrow) [125]. For tag removal the recognition sequence is usually located after the affinity tag but immediately before the first amino acid of the target protein. Using the pET28a expression vector system an N-terminal hexa-histidin-tag (20 amino acids; sequence: MGSSHHHHHSSGLVPRG↓SH, cleavage site shown as arrow) can be attached to the target protein by in-frame cloning of the respective target gene into the vector by using appropriate restriction endonucleases.

For tag removal biotinylated Thrombin (Novagen, Darmstadt, Germany) was used. After complete proteolytic cleavage biotinylated Thrombin can be removed from the sample by using Streptavidin agarose (Merck Millipore, Darmstadt, Germany).

After Ni-NTA purification of the recombinant PpSB1-LOV protein the Ni-NTA elution buffer (2.6.3.2) was exchanged by dialysis to Ni-NTA equilibration buffer (2.6.3.2). The protein concentration was determined using the Bradford assay (2.9.3.5) and 1 Unit of Thrombin was

added per mg of His-tagged protein. Cleavage was performed at 21 °C for 16 - 18 hours. Subsequently, the cleaved His-tag was removed by passing the sample over a Ni-NTA gravity flow column. The target protein as well as biotinylated Thrombin do not bind to the column and are thus found in the flow-through fraction. Finally the sample is passed through the Streptavidin column to remove biotinylated Thrombin from the sample. In this step Thrombin binds to the resin and the untagged target protein will be found in the flow-through fraction. Subsequently, the buffer is again exchanged by dialysis (2.9.3.2) for dialysis/storage/NMR buffer (2.6.3.3).

2.9.3.4 Concentration of protein samples by ultrafiltration

The purified protein was concentrated by ultrafiltration using 0.5 ml or 20 ml centrifugal concentrator units with a 10 kDa cutoff (Millipore, Schwalbach, Germany; or Vivascience, Goettingen, Germany). Centrifugation was performed at 4 °C at 5000 rpm. Final protein concentrations were estimated using the Bradford assay (2.9.3.5).

2.9.3.5 Determination of protein concentrations [126]

Protein concentrations were estimated using the Bradford method [126] with bovine serum albumin as standard [126].

To determine protein concentrations 900 µl of Bradford reagent (2.6.7) are mixed with 100 µl of a protein sample of appropriate dilution. The mixture is incubated for 10 min at room temperature. As reference the buffer in which the protein was dissolved was used. The concentration of the resulting protein-dye complex was determined spectrophotometrically at by measuring the absorbance at 595 nm. Based on a calibration curve recorded using known concentrations of BSA, the protein content in the sample is estimated.

2.9.3.6 Estimation of native protein molecular weights

For estimation of native protein molecular weights size exclusion chromatography (SEC) was employed. Ideally, SEC separates biomolecules such as proteins according to their size (molecular weight) and shape. In this type of chromatography, a variety of materials can be used as a stationary phase but generally polyacrylamide, dextran or agarose is used. Typically, aqueous buffer solutions that mimic physiological conditions are used as a mobile phase. Small molecules are trapped in the pores of the stationary phase and thus take longer to pass

through the column whereas larger molecules travel faster through the pores and elute earlier [127].

2.9.3.6.1 Protein molecular weight estimation by size exclusion chromatography (SEC) using a (high-pressure) HPLC system

HPLC-based size-exclusion chromatography experiments were performed on a LC-2010A HT HPLC system (Shimadzu, Corporation, Kyoto, Japan), equipped with a photodiode array (PDA) detector by employing a BioSep-SEC-S3000 HPLC column in the dimension 300/7.8 (Phenomenex, Aschaffenburg, Germany). All HPLC-SEC analysis were performed under the same conditions as the NMR measurement. Thus, a buffer similar to the one used in NMR experiments (HPLC-SEC running buffer for proteins, 2.6.6) heated to 40 °C was used as eluent under isocratic conditions (flow-rate: 1 ml/min). Likewise, the SEC column mounted within the HPLC column oven was heated to 40 °C. The elution of the protein was followed by UV detection at 220 nm and 280 nm.

In order to evaluate the influence of photo-activation on the quaternary structure of PpSB1-LOV SEC experiments were performed for the protein equilibrated in the dark and after photo-activation. For SEC runs to be carried out in the dark, all light exposed components of the HPLC instrument such as sample injector and outside tubings were wrapped in aluminium foil. The protein sample was pre-equilibrated overnight in the dark at 37 °C. For evaluation of the proteins light-state quaternary structure the protein sample was pre-illuminated for 1 min using a handheld V8-LED-Lenser (Zweibrüder Optoelectronics, Solingen, Germany) and immediately injected in to the HPLC system.

Hereby, the photodiode array detector of the employed HPLC system enables the recording of UV/Vis spectra of the respective protein elution peak. In this way it is possible to verify complete on-column population of the PpSB1-LOV dark- and light-state by visual inspection of the UV/Vis spectra of the protein elution peak.

2.9.3.7 HPLC-SEC analysis of peptide samples

To verify peptide purity HPLC-SEC experiments were performed on a LC-2010A HT HPLC system (Shimadzu, Corporation, Kyoto, Japan) by using a BioSep-SEC-S2000 HPLC-SEC column in the dimension 300/7.8 (Phenomenex, Aschaffenburg, Germany). HPLC-SEC

running buffer for peptides (2.6.6) was used as eluent under isocratic conditions (1 ml/min). The elution of the peptide was followed by UV detection at 220 nm and 280 nm.

2.9.3.8 Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) [118]

The purity of purified PpSB1-LOV protein preparations was assessed under denaturing conditions in an discontinuous gel system as described by Laemmli [118]. For protein denaturation the anionic detergent SDS was used. This results in linearization of the proteins polypeptide chain, whereas one SDS molecule binds to two amino acid residues. SDS binding gives the protein molecule a negative charge resulting in a defined mass to charge ratio. During electrophoresis protein molecules are separated according to their molecular weight as they migrate through the polyacrylamide matrix of the gel. Because of the relation between mass and charge, smaller molecule will travel faster through the polyacrylamide matrix of the SDS-PAGE gel.

For preparation of SDS-PAGE gels a discontinuous buffer system (2.6.4) was used. Stacking and separation gels were prepared as follows:

Table 9: Constituents and preparation of a 15% SDS-Gel

	separation gel	stacking gel
separation gel buffer (2.6.4)	2.5 ml	-
stacking gel buffer (2.6.4)	-	1.3 ml
Acrylamid / bisacrylamide	5.0 ml	0.83 ml
dH₂O	2.4 ml	2.8 ml
10 % (w/v) SDS	0.1 ml	0.05 ml
10 % /w/v) APS	0.1 ml	0.05 ml
TEMED	0.01 ml	0.01 ml

Protein samples were mixed with 4x protein loading buffer (2.6.4) and subsequently denatured at 100 °C for 5 min. 15 µl of sample were loaded per gel lane. Electrophoresis was performed in 1x SDS running buffer (2.6.4) using a vertical Mini-Protean II Dual Slap Cell system (Biorad, München, Germany). Gels were run at a constant voltage of 100 V. For

estimation of denatured protein molecular weights a standard mixture of proteins with known molecular weight (2.6.4) was used. SDS-PAGE gels were stained using a colloidal coomassie gel staining solution (2.6.4). Finally, the gels were photographed using a Stella gel documentation system (Raytest, Straubenhardt, Germany).

2.9.3.9 HPLC analysis of LOV protein flavin content and flavin species distribution

Depending on the expression conditions bacterial LOV proteins are known to bind any of the three naturally occurring flavin species, i.e. FAD, FMN and Riboflavin to a different degree [49]. This results in chromophore heterogeneity and thus probably structural heterogeneity of the corresponding LOV protein preparation. Therefore, it is important to verify the flavin species distribution of LOV protein preparations obtained under different expression conditions. For HPLC analysis, the protein bound chromophore was released by heat denaturation (99 °C, 20 min). The precipitate is removed via centrifugation (4 °C, 14000 rpm). Subsequently, the flavin chromophore containing supernatant is passed through a microcon YM3 (3 kDa cutoff) centrifugal concentrator unit (Millipore, Schwalbach, Germany) to remove residual protein contaminations. The resulting filtrate is then analyzed by HPLC by injecting 10 µl of the sample into a HPLC system. For separation of the different flavin species a reverse-phase C18- HPLC column in the dimension 250/4 equipped with an 8/4 pre-column filled with the same material (Macherey&Nagel, Düren, Germany) was used. As eluents 50 mM ammonium acetate, pH 6.0 (eluent B) and 70 % (v/v) acetonitrile in B (eluent A) were employed. The different flavin species (FAD, FMN and Riboflavin) elute according to their hydrophobicity i.e. more polar molecules will travel faster with the solvent through the column (FAD, FMN) and thus elute earlier than more unpolar ones (Riboflavin) [128]. Typical retention times for FAD, FMN and Riboflavin were as follows. FAD: 9.97 min, FMN: 10.75 min and Riboflavin: 12.94 min. Flavin species identification was achieved by using the respective authentic compounds as standards. The relative distribution of the different flavin species in the analyzed LOV protein preparation was estimated by comparing the corresponding elution peak areas.

2.9.3.10 Matrix Assisted Laser Desorption/Ionization-Time Of Flight Mass Spectrometry (MALDI-TOF MS) molecular mass analysis

The MALDI-TOF MS techniques was developed in 1988 by Prof. Hillenkamp [129]. It can be used to determine the molecular mass of proteins and small organic compounds. For MALDI-TOF MS analysis a solid matrix and the sample are mixed in an appropriate solvent, i.e. aqueous buffer. The mixture is subsequently dried on the MALDI plate, whereas different matrix materials can be used for different applications. For ionization a laser flash is used which results in effective ionization of the matrix molecules. The sample molecules are consequently ionized by proton transfer from the matrix and thus become mobile. The resulting sample molecule ions are accelerated in an electric field and are then analyzed in the MALDI-coupled time-of-flight (TOF) mass spectrometer. In the mass spectrometer these ions are separated according to their mass (m)-to-charge (z) ratios (m/z). The separated ions are then detected by a secondary electron multiplier tube.

In order to verify complete isotopic labeling of the LOV protein preparations as well as effective His-tag removal MALDI-TOF MS was used. All analyses were performed by Norbert Dickmann at the Max-Planck-Institute for Bioinorganic Chemistry using an ABI Voyager DE Pro MALDI-TOF mass-spectrometer. As matrix sinapinic acid (Sina) was employed.

2.9.4 Spectroscopic methods

2.9.4.1 Circular dichroism (CD) spectroscopy

Far-UV CD spectroscopy was used to investigate the secondary structure content of peptide samples. CD is observed when optically active matter (here the polypeptide backbone of peptide and protein molecules) absorbs left- and right-handed circular polarized light differently.

CD spectroscopic analysis of peptide unfolding/refolding

Peptide concentrations used for CD studies were approx. 60 μ M. To monitor the temperature dependent unfolding of the PpSB1-LOV and PpSB2-LOV J α -peptides, far-UV CD spectra were recorded sequentially while the temperature of the sample was raised from 1 $^{\circ}$ C to 65 $^{\circ}$ C using a constant ramp-rate of 1 $^{\circ}$ C / minute. Refolding of the respective peptide was followed

in a similar manner by cooling the completely unfolded peptide sample to 1 °C using the same ramp-rate.

α -peptide CD spectra were analyzed assuming the presence of only two pure components (random coil and α -helix) using the complex constraint analysis (CCA+) tool [49, 130]. Melting temperatures were obtained from the plot of the ellipticity at 222 nm ($\Theta_{222\text{nm}}$) versus the temperature by fitting the experimental data to equation 4. The same fit was used to estimate the maximal and minimal ellipticity values for the fully folded peptides $[\Theta]_{\text{folded}}$ and the completely unfolded peptides $[\Theta]_{\text{unfolded}}$. Those latter values were used to quantitatively describe α -helix melting in terms of the α -helical fraction ($f\alpha$), under the premise of a two-state transition (equation 5).

$$\theta_{222\text{nm}}(T) = \frac{[\theta]_{\text{folded}} - [\theta]_{\text{unfolded}}}{1 + e^{(T - T_m)/dx}} + [\theta]_{\text{unfolded}} \quad (4)$$

$$f\alpha = \frac{[\theta] - [\theta]_{\text{unfolded}}}{[\theta]_{\text{folded}} - [\theta]_{\text{unfolded}}} \quad (5)$$

Thermodynamic parameters were obtained by fitting the calculated equilibrium α -helix fraction ($f\alpha$) to equation 6.

$$\ln(K_{\text{eq}}) = \ln\left(\frac{f\alpha}{1-f\alpha}\right) = \frac{-\Delta H}{R} \left(\frac{1}{T}\right) + \frac{\Delta S}{R} \quad (6)$$

From this fit, values for the enthalpy ΔH and the entropy ΔS were derived. The relation $\Delta G = \Delta H - T\Delta S$ was used to obtain the Gibbs free energy of unfolding (ΔG) at a given temperature.

2.9.4.2 UV/Vis spectrophotometry

2.9.4.2.1 Measurement of the LOV protein photocycle associated spectral changes

All UV/Vis absorbance measurements were carried out under dim-red safety light. Light-dependent absorption changes in the UV/Vis region (200 - 600 nm) were recorded using a temperature-controlled Beckmann UV650 spectrophotometer (Beckmann Coulter, Fullerton, CA, USA). Protein samples were either diluted in 10 mM sodium-phosphate buffer supplemented with 10 mM NaCl (pH 8.0) or protein dialysis/storage NMR buffer (2.6.3.3). The same buffer was used as a reference. LOV protein light-state UV/Vis spectra were recorded after illuminating the dark-equilibrated sample for 30 seconds using a blue-light emitting Led-Lensers® V8 lamp (Zweibrüder Optoelectronics, Solingen, Germany). Dark-state recovery was measured from illuminated samples by monitoring the absorption-recovery at 485 nm. All dark-recovery kinetic measurements were at least carried out twice using two independent protein preparations. Dark recovery times traces were fitted using a single- or double exponential decay function by employing Origin 7G (OriginLab Corporation, Northhampton, MA, USA).

2.9.4.2.2 Estimation of the apo-protein content of LOV-protein preparations

Depending on the LOV-protein preparation, it is possible, that not all LOV-protein molecules contain a FMN chromophore. The presence of large amounts of such apo-protein molecules will result in unwanted structural heterogeneity in the preparation. This problem can be particularly severe when recombinant LOV-protein biosynthesis outbalances the endogenous FMN biosynthetic capabilities of the employed *E. coli* expression strain.

Therefore, UV/Vis spectrophotometry was used to estimate the amount of apo-protein present in all used LOV-protein preparations. Hereby, the concentration of the bound flavin chromophore was estimated by determining the absorbance of the LOV-protein solution at 447 nm. Absorbance in this wavelength range is solely dependent on the bound FMN chromophore. Using the molar extinction coefficient of FMN at 447 nm ($\epsilon_{\text{FMN}447\text{nm}} = 12500 \text{ M}^{-1} \text{ cm}^{-1}$, [59]) the absolute FMN concentration (c_{FMN} , in M) in the sample can be derived by using the Lambert-Beer relation (equation 7).

$$E_{\lambda} = \varepsilon_{\lambda} c d$$

(7)

Where E_{λ} = absorption at a given wavelength λ , ε_{λ} = molar extinction coefficient at the same wavelength (in $M^{-1} \text{ cm}^{-1}$), c = concentration of the molecule (M) and d = optical pathway of the cuvette (cm)

Since all flavins also absorb light in the UV region of the spectrum (i.e. at the same wavelength were aromatic protein amino acids absorb), the protein concentration in the sample cannot simply be estimated from the measurement of the protein absorbance at 280 nm. Therefore, it was necessary to subtract the contribution of the protein-bound FMN molecules at 280 nm from the protein absorption at 280 nm. Using the molar extinction coefficient of FMN at 280 nm ($\varepsilon_{\text{FMN}280\text{nm}} = 20300 \text{ M}^{-1} \text{ cm}^{-1}$) and the above calculated FMN concentration in the sample, the absorbance of FMN at 280 nm ($E_{\text{FMN}280\text{nm}}$) can be derived by using equation 7. This theoretical absorbance is subtracted from the measured absorbance of the sample at 280 nm. The resulting absorbance value should now solely derive from the apo-protein present in the sample. For estimation of the protein concentration (c_{Protein} , in M) of the sample the theoretical extinction coefficient ($\varepsilon_{\text{Protein}280\text{nm}}$) of the respective protein at 280 nm is used. Theoretical LOV-protein extinction coefficients were computed using the ExPASy-ProtParam tool (<http://web.expasy.org/protparam/>) [71].

The percentage of apo-protein in the sample is derived using equation 8.

$$\text{apo-protein} [\%] = \left(100 - \left(\frac{c_{\text{FMN}}}{c_{\text{Protein}}} \times 100 \right) \right)$$

(8)

2.10 Bioinformatic analysis

Protein sequences with significant similarity to the two short LOV proteins of *P. putida* KT2440, were obtained from the NCBI database by using the PSI-BLAST tool [131] employing the PpSB1-LOV and PpSB2-LOV amino acid sequences as query. To verify coverage of all *Pseudomonadaceae* genome sequences, the same search was conducted using

the NCBI genomic group BLAST tool (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) restricting the search to the group of *Pseudomonadaceae*. Domain content analysis were performed using the SMART tool (<http://smart.embl-heidelberg.de>) [132]. Amino acid sequence alignments were generated using the T-COFFEE alignment tool (http://www.igs.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi) [133]. Alignments were visualized and edited manually using GeneDoc [131]. The PCOILS webserver (<http://toolkit.tuebingen.mpg.de/pcoils>) [134] was used to analyse J α -helix sequences for the presence of a coiled-coil heptad repeat pattern. Charges and average hydrophathies (according to the Kyte and Doolittle scale [135]) were inferred from sequence alignments using custom-made Perl scripts. Mapping of NMR data to PDB X-ray coordinates was achieved by using custom-made Perl scripts. Helical-wheel and coiled-coil helical wheel analyses were performed using the WheelApp Java applet (<http://cti.itc.virginia.edu/~cmg/Demo/wheel/wheelApp.html>) and the DrawCoil 1.0 tool provided by the Grigoryan Lab (Dartmouth College, Hanover, USA; www.grigoryanlab.org/drawcoil).

2.11 NMR spectroscopy

2.11.1 Peptide NMR measurements – sample preparation

The two peptide samples (PpSB1-J α and PpSB2-J α) were dissolved in 100 mM sodium phosphate buffer pH 6.3 supplemented with 50 mM NaCl before [D₂]TFE was added to yield a final TFE concentration of 10 % (v/v) (PpSB1-J α) or 25 % (v/v) [D₂]TFE (PpSB2-J α), respectively. Peptide concentrations used for NMR studies were approx. 1 mM. Additionally, all samples contained 10 % (v/v) D₂O. Samples were transferred to a 5 mm NMR shigemi tube (Sigma-Aldrich, Steinheim, Germany) and placed into the NMR instrument.

2.11.2 LOV-protein NMR measurements – sample preparation

Isotopically labeled protein samples were prepared as described in 2.9.1.5.3. Protein samples used to perform all NMR experiments contained 0.5 – 1 mM (¹³C-¹⁵N) or (¹⁵N) isotopically labeled PpSB1-LOV protein diluted in dialysis/storage NMR buffer (2.6.3.3) supplemented with 10 % (v/v) D₂O and 0.05 % (v/v) sodium azide. Before the measurements the protein sample was transferred in the dark to a 5 mm NMR shigemi tube (Sigma-Aldrich, Steinheim,

Germany) wrapped in aluminium foil. For dark-state measurements, this sample was transferred to the NMR instrument in the dark by covering the shigemi tube with aluminium foil and by keeping all lights off. To perform light-state measurements, the sample was pre-illuminated in the shigemi-tube using white-light until the yellow colour of the sample disappeared. This phenomenon is due to the fact that the LOV protein in its light-state does not absorb blue light anymore (no absorbance at 450 nm, see chapter 1.2.4 of the introduction). Illumination during NMR experiments was achieved by using blue-light emitting LED (470 nm) which was coupled into a fibre optical cable via a collimator lens. The end of the fibre optic was placed into the shigemi tube inside the NMR magnet. In this way it is possible to continuously illuminate the sample during the measurement.

2.11.3 Data Collection

All final NMR experiments were carried out at 40 °C on either a Varian NMR instrument operating at a ^1H Larmor frequency of 600 MHz or an Varian NMR instrument operating at a ^1H Larmor frequency of 900 MHz. The NMR spectrometer was equipped with cryogenically cooled ^1H - ^{13}C - ^{15}N triple resonance probe with Z-axis pulse-field-gradient (PFG) or room temperature PFG triple resonance probes at proton frequency of 900 MHz or 600 MHz. 2D ^1H - ^1H NOESY (Nuclear Overhauser Effect Spectroscopy) with 350 ms (PpSB1- $J\alpha$) or 450 ms (PpSB2- $J\alpha$) mixing time and 2D ^1H - ^1H TOCSY (Total Correlated Spectroscopy) spectra with 60 ms mixing time were recorded for both peptide samples. Standard pulse-sequences were used for all experiments. NMR acquisition and data processing parameter of the performed NMR experiments are shown in Table A.1 in the appendix.

2.11.4 Processing

After spectra acquisition, the time domain data was processed and converted into the frequency domain with the aid of the NMRPipe software [112]. NMR data was processed using NMRPipe/NMRDraw [112] and evaluated with the CcpNmr Analysis software [106].

2.11.5 Resonance assignment for $J\alpha$ -peptide samples

For the peptide resonance assignment ^1H - ^1H TOCSY [136] and ^1H - ^1H NOESY [137] experiments were performed using unlabeled peptide samples. For resonance assignment one typically starts with a residue that ideally occurs only once in the peptide. Assignment is then achieved by comparison of the experimentally observed chemical shift values for a given

amino acid to tabulated values [138]. This allows grouping of all observed signals into a spin system corresponding to the amino acid type. Once all signals have been grouped into the spin system, a NOESY spectrum is used for sequential resonance assignment. The NOE is most important source for structure calculation. Typically, NOESY spectra of proteins contain thousands of cross peaks for atoms that are within approximately 6 Å of each other. Cross peak intensities scale by the inverse sixth power of the distance between the two nuclei. NOESY spectra are used to connect the amide proton of a given residue to the α -proton and amide proton of the previous and next residue.

The peak volume, as a measure of the NOESY cross peak intensity, is derived from integration of the corresponding peak. Peak volumes translate into distances according to equation 9.

$$V = c \frac{1}{d^6} \quad (9)$$

Where V is the volume of the NOESY peaks, c is a calibration constant and d is the spatial distance between the atoms involved. The structure of the protein or peptide can be calculated by a sufficiently complete set of such distance constraints (Fig. 8).

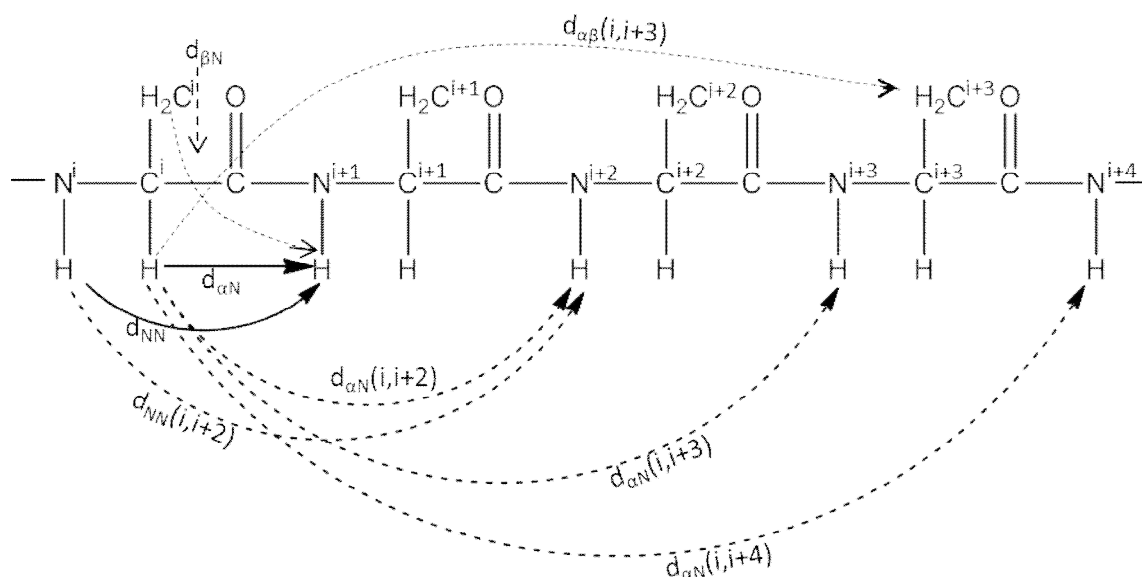


Figure 8: Sequential and medium-range ^1H - ^1H distances in polypeptide chains [138]. Solid arrows indicate adjacent amino acids and sequential ^1H - ^1H distances. Dotted and broken arrows visualize non-sequential and medium/long range ^1H - ^1H distances useable for sequential resonance assignment (Fig taken from “NMR of Proteins and Nucleic Acids” Wuthrich, p118, (1986)).

2.11.6 Sequential assignment of protein resonances

2.11.6.1 Backbone resonance assignment of PpSB1-LOV

The assignment of backbone and side-chain resonances to individual nuclei in the protein is the first step for extracting any information at atomic resolution from NMR data. To obtain a backbone resonance assignment the following 2D and 3D NMR experiments were measured:

2D ^1H - ^{15}N HSQC [139]

3D HNCO [140]

3D HNCA [140]

3D HN(CO)CA [141]

3D HN(CA)CO [142]

3D ^1H - ^1H - ^{15}N -NOESY-HSQC [143]

Triple resonance backbone experiments are used to correlate nuclei of a specific residue i with nuclei of the neighboring amino acid $i-1$.

A schematic representation of the NMR experiments used to obtain the backbone resonance assignment for PpSB1-LOV are shown in Fig. 9, panel A. Side-chain resonance assignment of methyl-containing amino acids for the PpSB1-LOV light-state was achieved by using the 3D HCCH-TOCSY experiment (panel B) and 3D ^{13}C -NOESY experiments.

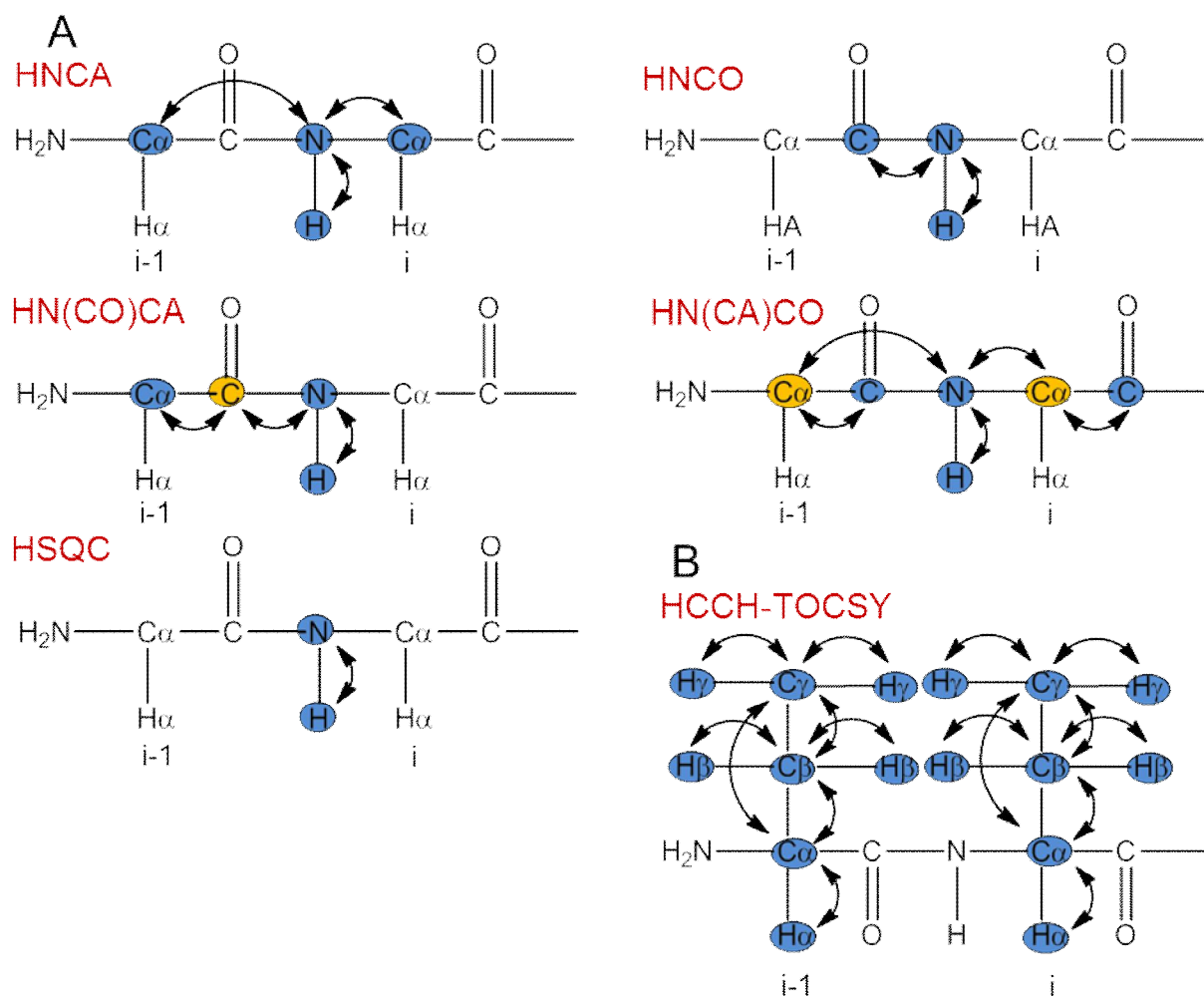


Figure 9: Schematic overview of the heteronuclear NMR experiments used in this study for the assignment of backbone and side-chain resonances. Blue circles indicate the correlations that are recorded in the given experiment; yellow circles represent nuclei for which no chemical shift is recorded, but that are used for coherence transfer. Experiments used for the backbone resonance assignment are shown in panel A. Side-chain assignment of methyl groups was achieved by using the 3D HCCH-TOCSY experiment (panel B).

2.11.7 Chemical shift perturbation

Chemical shift differences between light- and dark-state 3D HNCO spectra were quantified by the minimum chemical shift difference method [144] according to equation 10.

$$\Delta\delta_{\min} = [\Delta\delta(^1\text{H})^2 + 0.17\Delta\delta(^{15}\text{N})^2 + 0.39\Delta\delta(^{13}\text{C})^2]^{1/2} \quad (10)$$

Where $\Delta\delta(^1\text{H})$ represents the chemical shift difference between the respective proton resonance in the light-state and the dark-state, $\Delta\delta(^{15}\text{N})$ and $\Delta\delta(^{13}\text{C})$ are the respective nitrogen and carbon chemical shift differences obtained from the respective light- and dark-state 3D HNCO spectra. In order to normalize the magnitude of the ^{15}N and ^{13}C chemical shift change to that of ^1H chemical shift changes, scaling factors of $\alpha_{\text{N}} = 0.17$ and $\alpha_{\text{CO}} = 0.39$ were used [144]. In this way the residue-wise (backbone) chemical shift perturbation resulting from PpSB1-LOV photo-activation can be quantified.

2.11.8 NMR structure calculation

In order to derive structural information from NMR experiments distance restraints have to be obtained for the different NMR active nuclei of the sample. This structural information is accessible from Nuclear Overhauser Effect (NOE) experiments. Hereby, the structure of i.e. the peptide is determined by distance restraints derived from cross peaks in a NOESY spectrum. The NOE is result of cross relaxation via dipole/dipole interactions between two NMR active nuclei (i.e. protons in case of the unlabeled peptide samples). The NOE between two nuclei is proportional to the $1/r^6$, where r is the distance between the two nuclei [138]. Usually structures can be derived from a collection of such signals which define distance constraints between i.e. the hydrogen atoms along the polypeptide chain.

For peptide structure calculations, ^1H - ^1H NOESY spectra were used. NOESY cross peak assignment and structure calculations were performed using the combination of the programs Aria 2.3.1 [105] and CNS 1.21 (including the Aria patchset) with the PARALLHDG v5.3 forcefield. All MD parameters were used in default configuration but the numbers of steps were increased to 8,000 during refinement and 20,000 in each cooling phase. The 10 least energy structures out of 250 calculated were further refined in explicit water as a final step in the Aria procedure. Geometry parameters of the derived structural models and the corresponding secondary structure distribution were analyzed and visualized using the programs MOLMOL [145], PyMOL [113], WHATIF [146] and MolProbity [111].

2.11.9 Hydrogen-deuterium (H/D) exchange

For H/D exchange experiments (see 1.3.5.3 for details) protein samples were concentrated to 300 μl possessing a final concentration of 0.5 mM. The same buffer as in all NMR experiments (2.6.3.3) was used. The protein was snap frozen in liquid nitrogen and dried by

lyophilization. Immediately before the first ^1H - ^{15}N HSQC experiment D_2O (300 μl , 99.8%) was added to the lyophilized sample and the protein was resuspended by gently pipetting. Subsequently a series of ^1H - ^{15}N HSQC spectra were recorded every 7.5 minutes. The first ^1H - ^{15}N -HSQC data points could be obtained after roughly 5 minutes. The last spectrum was recorded after 470 minutes in case of the dark-state measurement and after 110 minutes in case of light-state measurement. Total measuring time for each experiment was about 7 min. For light-state H/D exchange experiments a fresh sample, prepared as described above, was used. The sample was kept in the dark during all the preparative steps and was illuminated with blue light after addition D_2O .

Data was analyzed with CcpNmr, relative cross peak intensities were extracted and fitted using a single exponential decay function.

2.11.10 Heteronuclear ^1H - ^{15}N NOE (hetNOE), longitudinal and transverse relaxation experiments

All spin relaxation experiments (see 1.3.5.3 for details) for the PpSB1-LOV protein in the dark- and light-state were conducted at two static magnetic field strengths (600 and 800 MHz). ^1H - ^{15}N -NOE, ^{15}N longitudinal (T1) and transverse (T2) relaxation experiments were measured at 40 °C using ^{15}N - labeled samples. ^1H - ^{15}N -NOE-TROSY spectra were acquired with 1.8 seconds proton saturation. For T1 measurements, relaxation delays of 10, 20, 60, 120, 350, 500, 700, 900, 1200, 1800 ms were recorded. For T2 measurements, data was acquired with delays of 10, 30, 50, 70, 90, 110, 130 and 150 ms. ^1H - ^{15}N NOEs were collected with and without proton saturation. Data of T1 and T2 relaxation experiments were fitted to a mono-exponential decay function using CcpNmr to obtain corresponding T1 and T2 decay time constants. Those can be converted into rates by $R1 = 1 / T1$ and $R2 = 1 / T2$.

2.11.11 Deposition of NMR assignments

The PpSB1-J α and PpSB2-J α resonance assignments have been deposited at the Biological Magnetic Resonance Bank (BMRB, <http://www.bmrb.wisc.edu/>) under accession numbers 18159 and 18160. The assignment information for the two peptides is listed in Table A.3 and Table A.4 in the appendix. The atomic coordinates of the resulting PpSB1-J α and PpSB2-J α structures have been deposited at the RCSB protein data bank under accession numbers 2YON and 2YOM. The backbone and side-chain assignments for the PpSB1-LOV dark- and

light-state have not been deposited. The respective information is summarized in Table A.5 and Table A.6 in the appendix.

Chapter 3: Results

The presented thesis focuses on the two short LOV proteins PpSB1-LOV and PpSB2-LOV of *P. putida* KT2440, which as simplistic full-length prototypic photoreceptor proteins yet lack a fused effector domain. Methodically solution NMR spectroscopy as well as complimentary biochemical and biophysical methods have been used to address structural conservation and photo-activation of PpSB1-LOV and PpSB2-LOV.

In general, the presented work is organized in two parts:

- i) The **role and function of the PpSB1-LOV and PpSB2-LOV C-terminal J α -extensions** was studied by truncation experiments in that variants of PpSB1-LOV and PpSB2-LOV that lack the respective N- and C- terminal extensions were generated. Moreover, NMR spectroscopy and CD spectroscopy was used to probe the folding, structure and stability of the isolated C-terminal (J α) extensions of both PpSB1-LOV and PpSB2-LOV.
(Reprinted with permission from (Rani, R., et al., Conservation of Dark Recovery Kinetic Parameters and Structural Features in the *Pseudomonadaceae* "Short" Light, Oxygen, Voltage (LOV) Protein Family: Implications for the Design of LOV-Based Optogenetic Tools. *Biochemistry*, 2013. **52**: p. 4460-73) copyright (2013) American Chemical Society)
- ii) The **conformational changes associated with PpSB1-LOV photo-activation** were studied by means of solution NMR techniques. The presented data include optimization of expression conditions for isotopic labeling, optimization of NMR buffer and measuring conditions as well final data sets obtained for complete backbone resonance assignment of both the dark- and light-state of PpSB1-LOV. Subsequently, H/D exchange experiments as well as relaxation experiments (T1, T2, and hetNOE) were conducted to assess dynamic properties of the protein in both states.

3.1 Conservation of auxiliary structural features in the two *P. putida* KT2240 short LOV proteins

Bacterial LOV proteins commonly occur as multi-domain proteins where the LOV sensor domain is bound to a variety of different effector domains, regulating a number of different cellular functions. As outlined in the introduction in some bacteria as well as in fungi LOV proteins exist, which lack a fused effector domain. Examples include the so-called short LOV proteins PpSB1-LOV and PpSB2-LOV of *P. putida* KT2440 [48, 49], which are the focus of this study, as well as architecturally related proteins such as VVD of *N. crassa* [54] and the RsLOV protein of *R. sphaeroides* [55, 147]. Instead of a fused sensor domain short LOV systems contain many times short N- and or C-terminal extensions which are commonly termed N-terminal cap (N-cap) and C-terminal J α -helix [48, 49, 54, 55, 147]. In bacteria those short LOV proteins are present in about 11-13 % of fully sequenced bacterial genomes [53]. The two *P. putida* short LOV proteins are highly similar with respect to their amino acid sequence (approx. 66 % identical amino acid positions), but display drastically different photocycle kinetic properties [49]. While their photochemistry has been studied in great detail [36, 56], little is known about their physiological role and the sensor system inherent photo-activation and signal-relay mechanism. Moreover, at the time this study was conceived and carried out the auxiliary structural segments (N-cap and J α) outside the conserved LOV core of PpSB1-LOV and PpSB2-LOV had not been truly structurally characterized and their role for function and folding of the full-length photoreceptor protein remained unclear. In the short LOV protein family those auxiliary structural elements appear unrelated in sequence, but were suggested to be at least in part α -helical [48, 54, 55]. For the two *P. putida* LOV proteins this hypothesis was based solely on bioinformatic analyses and CD spectroscopic data recorded for the full-length proteins [48, 49]. In the first part of the presented thesis, data regarding the structure, conservation and the functional role of the N- and C-terminal extensions for the folding of the full-length PpSB1-LOV and PpSB2-LOV proteins will be presented.

3.1.1 Cloning and expression of PpSB1-LOV and PpSB2-LOV truncated variants

In order to gain more insight into the structural role of the N-terminal cap (N-cap) and the C-terminal J α -helix present outside the LOV-core domain of the two *P. putida* short LOV-proteins, a set of truncated protein variants were generated using molecular biological means. In turn, the respectively truncated variants could then be studied by CD and NMR spectroscopy to elucidate the functional role and structure of the two structural elements for the folding and light-dependent signal-relay within the corresponding LOV protein.

The truncated PpSB1-LOV and PpSB2-LOV variants were amplified by PCR (2.9.2.5) from pET28a vectors containing the full-length PpSB1-LOV and PpSB2-LOV encoding genes [49] by using gene specific primers (Table 5, Materials and Methods 2.1.3), designed to allow directional cloning of the respective truncated variant by employing 5'- *Nde*I and 3'- *Xho*I restriction endonuclease sites. At the 3'- end of the respective gene fragment a TAA stop-codon was added. The truncated constructs consisted of the following protein segments: PpSB1- Δ Ncap (residues: 17 – 142), PpSB1- Δ J α (residues: 1 – 119), PpSB1- Δ Ncap Δ J α (residues: 17 – 119), PpSB2- Δ Ncap (residues: 17 – 148), PpSB2- Δ J α (residues: 1 – 119) and PpSB2- Δ Ncap Δ J α (residues: 17 – 119). Fig. 10 depicts a multiple sequence alignment of the respectively truncated constructs.

The resulting PCR products were digested using *Nde*I and *Xho*I (2.7) and were ligated with a respectively hydrolyzed pET28a vector (Novagen/Merck, Darmstadt, Germany) (2.9.2.4). Ligation reactions were transformed into *E. coli* DH5 α (2.9.1.4). Plasmid DNA of several transformands was purified (2.9.2.4.1), restricted using *Nde*I and *Xho*I (2.7) and visualized on an 2% agarose gel. Plasmids showing the desired restriction pattern were sent for DNA sequencing (2.9.2.6).

Plasmid-DNA of all sequence verified truncated constructs was transformed into *E. coli* BL21 (DE3) for expression (2.9.1.4). All constructs were expressed in *E. coli* in identical manner as the corresponding full-length proteins (2.9.1.5.1). Cells were lysed and insoluble material was separated by centrifugation (2.9.3.1.1). The resulting protein fractions (insoluble pellet and soluble protein containing supernatant) were separated by SDS-PAGE (2.9.3.8).

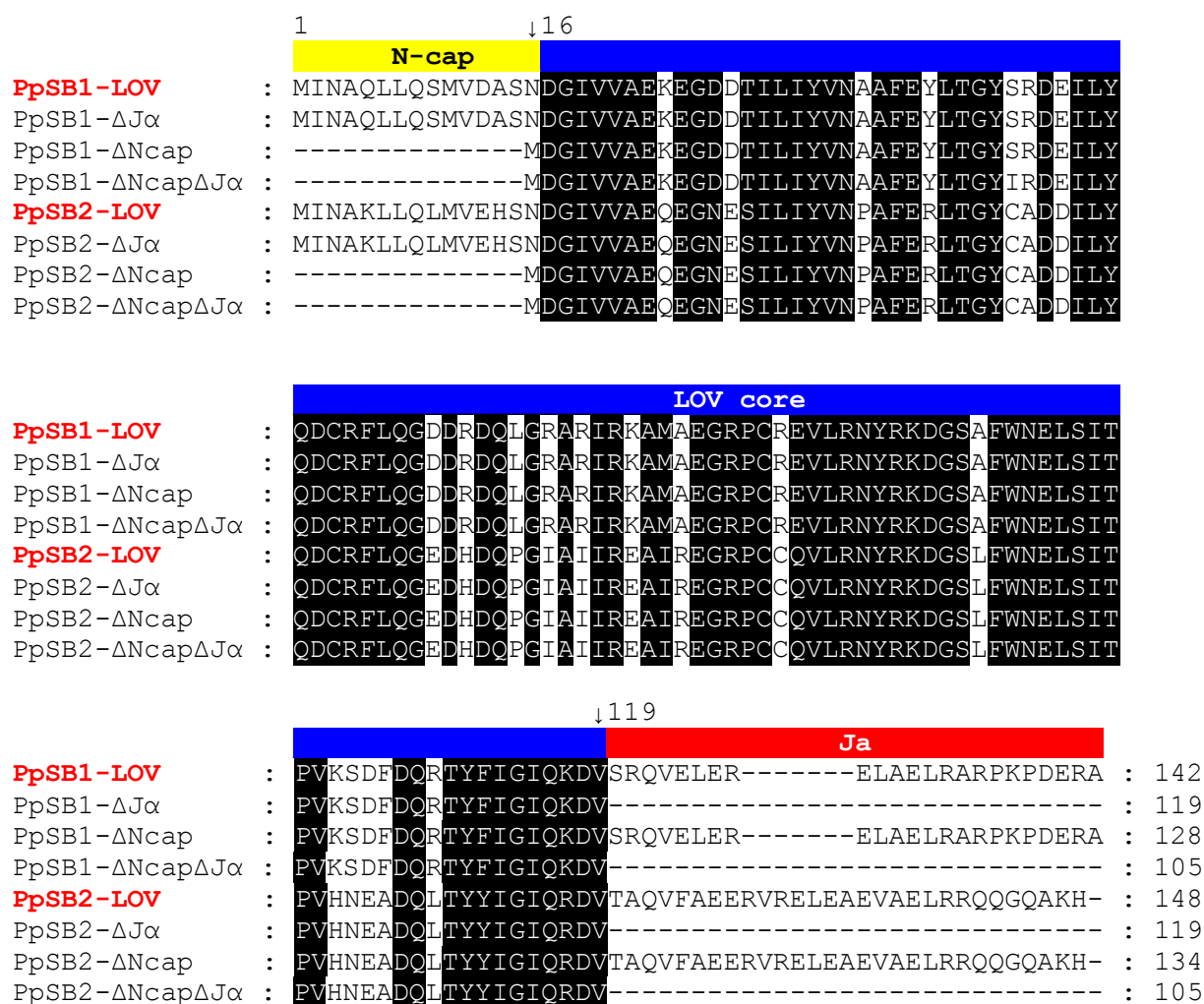


Figure 10: Multiple amino acid sequence alignment of the two wild-type full-length PpSB1-LOV, PpSB2-LOV proteins (highlighted bold in red) and the respective N-terminally and C-terminally truncated LOV-protein variants. Residues shaded in black are conserved in >90 % of all sequences. Structural features are shown in yellow: N-cap (residues 1 - 16), red: J α -helix (residues 119 - 142/148) and blue: LOV-core (residues 17 - 119) above the alignment. Positions where truncation was carried out are marked by an arrow.

Whereas both full-length proteins are present largely in the soluble fraction (solubility ≥ 50 %), removal of the N-cap and C-terminal J α -helix as well as of both structural elements, results in the accumulation of predominately insoluble aggregates of the respective protein construct.

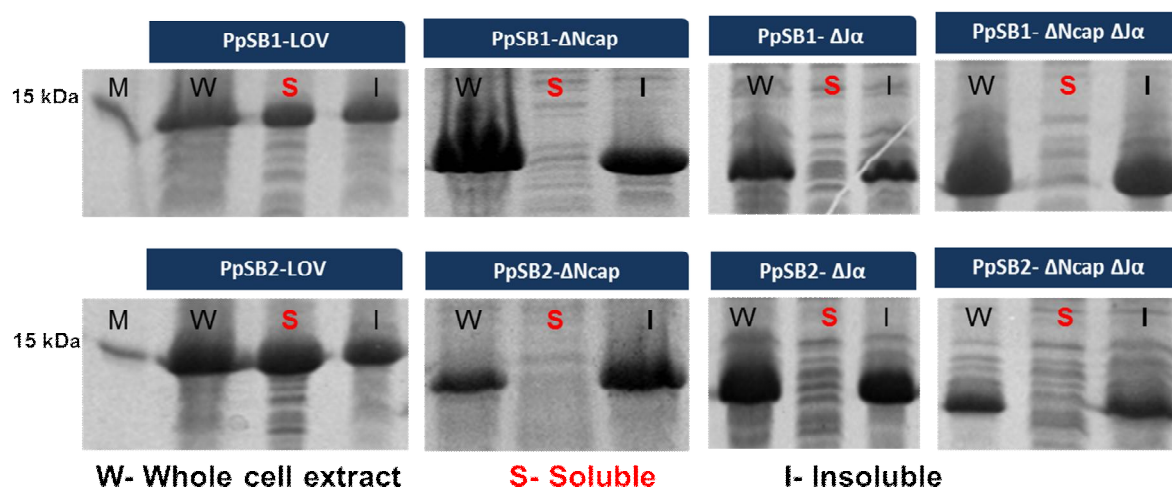


Figure 11: SDS-PAGE expression analysis. Recombinant proteins were expressed by growing *E. coli* BL21 (DE3), containing the appropriate expression plasmids (2.1.2), at 37 °C for 3 hours in autoinduction full-media (AI-FM, 2.9.1.5.1). Subsequently the temperature was reduced to 30 °C and cells were grown for additional 24 hours in the dark. After cell lysis by french press (2.9.3.1.1), unbroken cells and cell debris was separated from the soluble protein fraction by centrifugation. The insoluble pellet was resuspended in the initial volume of buffer and designated as insoluble fraction. The protein concentration of the soluble fraction was determined using the Bradford assay (2.9.3.5). For SDS-PAGE analysis approx. 20 μ g of soluble protein and the corresponding volume of the insoluble fraction was loaded per lane. Abbreviations W: whole cell extract, S: soluble fraction, I: insoluble fraction.

SDS-PAGE expression analysis of N- and C-terminally truncated PpSB1-LOV and PpSB2-LOV variants suggests that both structural elements, namely N-cap and J α -helix, contribute to proper folding and stability of the full-length photoreceptor protein during recombinant protein biosynthesis *in vivo*.

3.1.2 Bioinformatic analysis of the N- and C-terminal auxiliary sequence elements of PpSB1-LOV and PpSB2-LOV

As shown in the previous chapter, removal of both N- and C-terminal extensions from the full-length PpSB1-LOV and PpSB2-LOV proteins did not yield any SDS-PAGE detectable amounts of soluble protein (Fig. 11). Therefore, the truncated constructs could not be obtained in purified form for further biochemical/biophysical and structural studies. We thus initially carried out further bioinformatic analyses for the two auxiliary structural elements of both proteins. Fig. 12 depicts a multiple sequence alignment and helical wheel analysis of the

N-cap extension of various homologous short LOV proteins found in a number of *Pseudomonadaceae*.

Comparative sequence analyses reveal a high conservation of N-cap sequences in various *Pseudomonadaceae* short LOV proteins. Most sequences share 60 – 100% identical amino acid positions with the N-cap sequence of PpSB1-LOV and PpSB2-LOV. All sequences are rich in non-polar (hydrophobic) amino acids (Fig. 12) with the respective residues being localized around all sides of an ideal α -helix visualized by helical wheel plots for the two N-cap segments of PpSB1-LOV and PpSB2-LOV (Fig. 12, panel B,C).

In contrast to the highly conserved N-cap sequences, the C-terminal extension (J α -helix) is much more variable both in length and sequence. For dimeric photoreceptors such as *B. subtilis* YtvA, LOV-HKs and PAS chemosensors it was suggested that the sensor and effector domain are linked by a short α -helical linker that forms a parallel coiled-coil in the dimeric structure of the respective sensor proteins. In order to elucidate, whether such structural features might be conserved in *Pseudomonadaceae* short LOV proteins, we analyzed the respective protein sequences for the presence a canonical *abcdefg* heptad repeat pattern characteristic for coiled-coil structures (Fig. 13, panel A). For all sequences, the coiled-coil prediction tool PCOILS [134] predicts the presence of either two or three heptad repeats (Fig. 13, panel A). This is also apparent in the alignment where two distinct sequence length patterns can be identified. In a subset of sequences a conserved 7 amino acid insertion is present after the first heptad repeat accounting for two helical turns in structure. Moreover, the structural interaction pattern typical for coiled coils, featuring hydrophobic interactions via the residues *a - d* and salt-bridge contacts via the residues *e - g* of the heptad, is conserved in all sequences (Fig. 13, panel B and C).

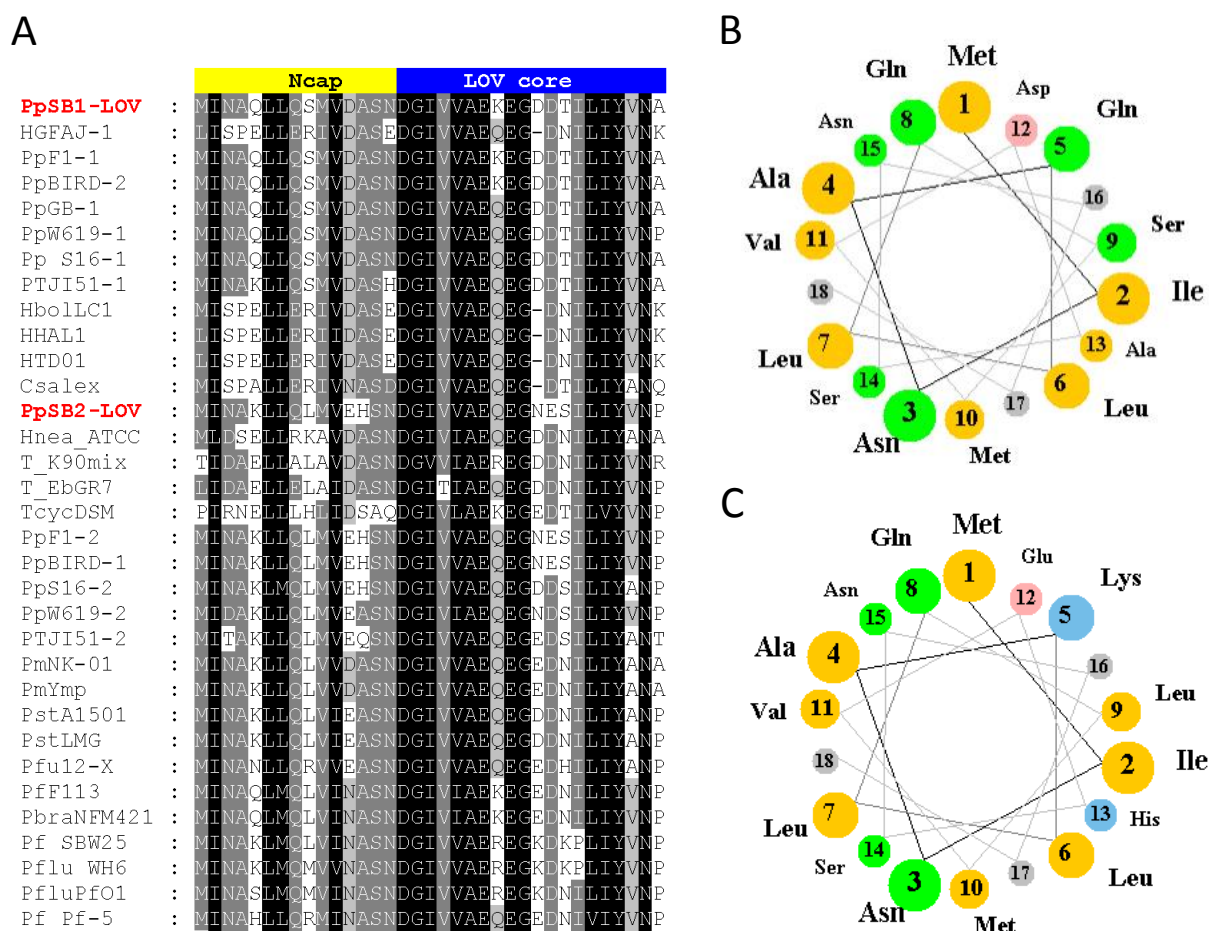


Figure 12: Multiple sequence alignment (A) and helical wheel analysis (B, C) of various *Pseudomonadaceae* short LOV protein N-cap amino acid sequences. To obtain the sequence alignment shown in (A) putative full-length short LOV sequences were obtained from the Genbank database (www.ncbi.nlm.nih.gov) by using the PSI-BLAST [148] tool. Sequence-IDs, organism names and accession numbers for the sequences shown in the alignment are listed in Table A.2 in the appendix. Sequences were aligned using ClustalW and truncated to show only the N-cap region (residue 1 – 16; highlighted in yellow above the alignment) and the first two β -strands of the conserved LOV core (in blue). (B) Helical wheel plot of the PpSB1-LOV N-cap sequence (C) Helical wheel analysis performed using the PpSB2-LOV N-cap sequence. Polar, uncharged amino acids are shown green, basic amino acids in blue, acidic amino acids are highlighted in salmon and non-polar ones are shown in orange. All helical wheel analysis was performed using the WheelApp Java applet provided by the University of Virginia (<http://cti.itc.virginia.edu/~cmg/Demo/wheel/wheelApp.html>).

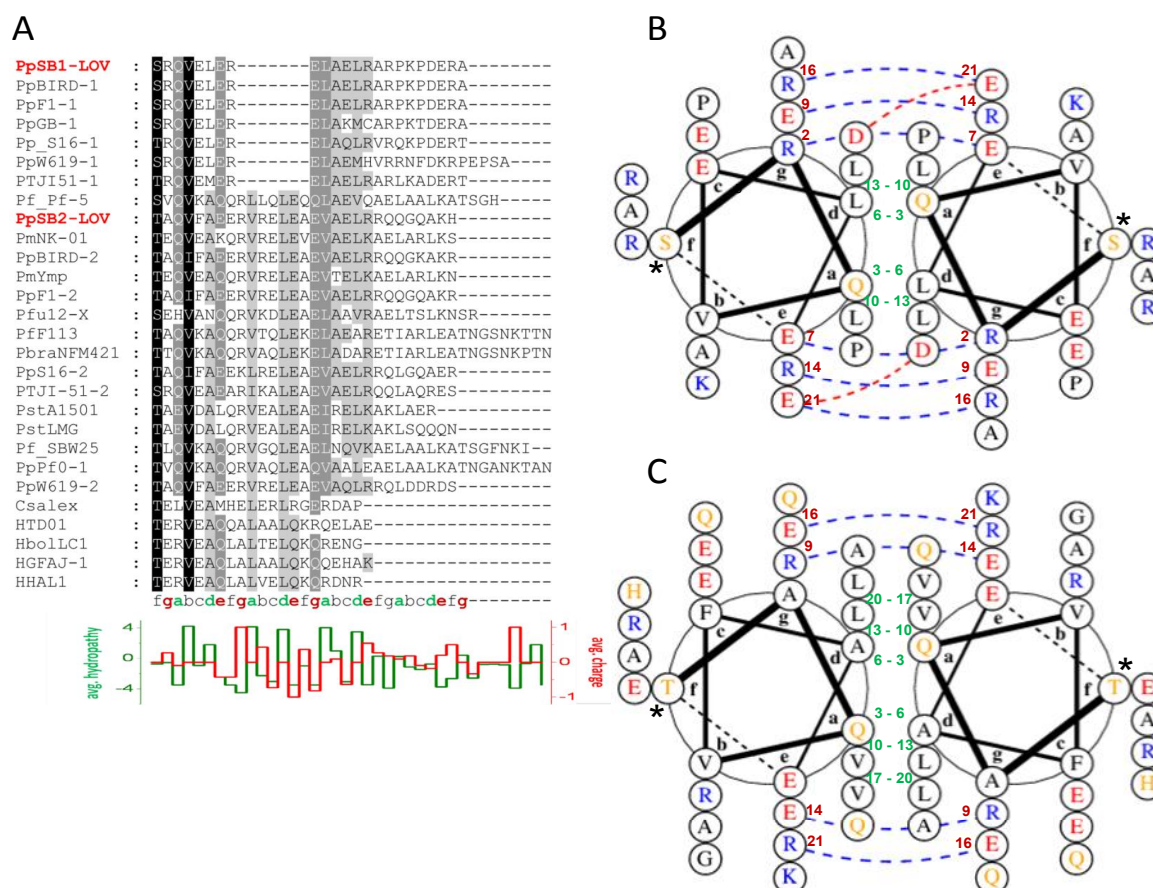


Figure 13: Multiple sequence alignment (A) and helical wheel (B, C) analysis of various *Pseudomonadaceae* short LOV protein $J\alpha$ -helix amino acid sequences. The alignment was truncated to show only the $J\alpha$ -helix region. Below the alignment average hydrophathies (in green) as well as average charges (in red) inferred from the alignment are plotted. Average residue wise hydrophathies were assigned according to the Kyte-Doolittle scale [135]. The PCOILS predicted heptad repeat pattern typical for coiled coils, featuring hydrophobic interactions via the residues *a* - *d* and salt-bridge contacts via the residues *e* - *g* of the heptad, is shown in the last row of the alignment. (B) Dimeric parallel coiled-coil helical wheel representation of the PpSB1-LOV $J\alpha$ -helix sequence. (C) Coiled-coil helical wheel representation of the PpSB2-LOV $J\alpha$ -helix sequence. Canonical coiled-coil heptad repeat residues enabling hydrophobic helix-helix packing are numbered in green. Residues forming the e-g salt-bridges of the coiled-coil are numbered in red. The starting amino acid of the $J\alpha$ -helix is marked by an asterisk. Unpolar residues are colored in black, basic amino acids in blue, acidic residues are shown in red and polar uncharged residues in orange. The diagrams shown in panel B and C were generated using the DrawCoil 1.0 tool provided by the Grigoryan Lab (Dartmouth College, Hanover, USA; www.grigoryanlab.org/drawcoil).

Bioinformatic analysis for a set of sequences of the *Pseudomonadaceae* short LOV family, suggest the presence of a well-defined N-cap as well as of C-terminal $J\alpha$ -helix elements prone to mediate subunit association via coiled-coil interactions, as suggested for dimeric multi-domain light (LOV) and chemosensory (PAS) systems.

3.1.3 Structural studies on the isolated PpSB1-LOV and PpSB2-LOV J α -peptides

3.1.3.1 Quality of the custom synthesized PpSB1-J α and PpSB2-J α peptides

To investigate the folding of the isolated C-terminal J α -extensions, the respective polypeptide segments of PpSB1-LOV and PpSB2-LOV were custom synthesized in N-acetylated (ACE) form (Peptide 2.0, Chantilly, VA, USA) (2.8) and studied by CD spectroscopy (2.9.4.1) and NMR techniques (2.11.1). The sequence of the two peptides is listed below:

PpSB1-J α : ACE-S₁₂₀RQVELERELAE LRARPKPDERA₁₄₂

PpSB2-J α : ACE-T₁₂₀AQVFAEERVRELEAEVAELRRQQGQAKH₁₄₈

To verify peptide purity, 0.2 - 0.4 mg/ml lyophilized powder of the PpSB1-J α and PpSB2-J α were dissolved in 200 mM sodium phosphate buffer (pH 7.5) supplemented with 10 mM NaCl and analyzed by HPLC-SEC (2.9.3.7) (Fig. 14).

PpSB1-J α was highly pure (data not shown) whereas the PpSB2-J α peptide preparation showed impurities of lower molecular weight (Fig. 14, A). Therefore the PpSB2-J α peptide preparation was further purified before CD and NMR spectroscopic measurements. Therefore, the peptide sample was dialyzed against 200 mM sodium phosphate buffer (pH 7.5) supplemented with 10 mM NaCl using a 1 kDa cut-off membrane (2.9.3.2). Subsequently, the thoroughly dialyzed sample was subjected to HPLC-SEC analysis to verify the removal of impurities (Fig. 14, panel B).

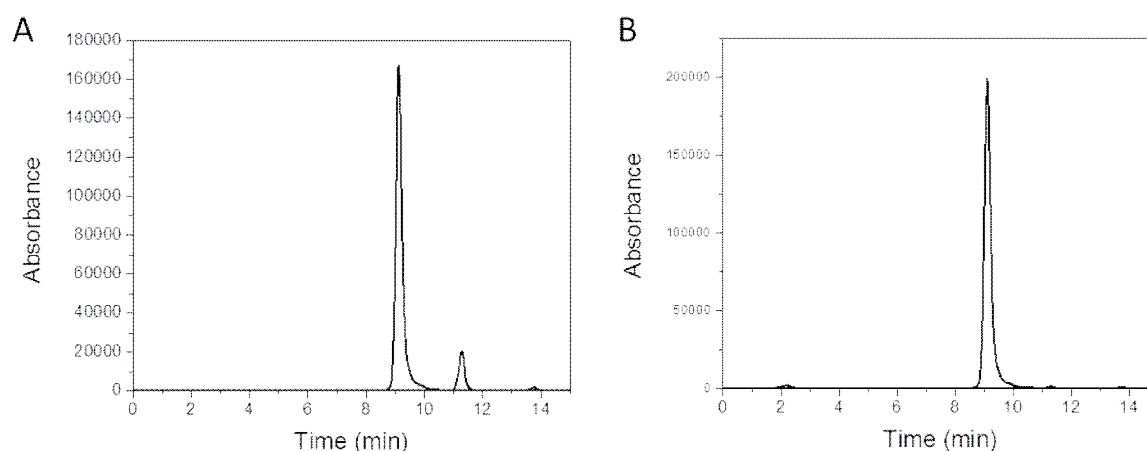


Figure 14: HPLC-SEC chromatogram of the custom synthesized PpSB2-J α peptide preparation before (A) and after (B) purification by dialysis. Peptide samples were separated by HPLC-SEC (2.9.3.7) using 200 mM sodium phosphate buffer (pH 7.5), supplemented with 10 mM NaCl as eluent.

After dialysis of the PpSB2-J α peptide preparation both samples were equally pure (> 98%) and were used for further studies. For NMR studies the peptide sample was dialyzed against 100 mM sodium phosphate (pH 6.3) supplemented with 50 mM NaCl.

3.1.3.2 CD spectroscopic studies reveal independent folding of the isolated PpSB1-J α and PpSB2-J α peptides

Coiled coils are widespread, many times independently folding [149], interaction motifs present in many structurally different proteins. Therefore, we asked the question whether the J α -helix extensions of PpSB1-LOV and PpSB2-LOV, that are predicted to form parallel dimeric coiled-coil structures (see results chapter 3.1.2), can fold independently from the rest of the protein. To study folding and secondary structure of both peptides we used far-UV CD spectroscopy (2.9.4.1). Both peptide samples were resuspended in 20 mM sodium phosphate buffer (pH 7.5), supplemented with 1 mM NaCl. Initially CD spectra were recorded at 20 °C. At this temperature clear spectral evidence for the presence of α -helical secondary structure could only be obtained for PpSB2-J α (data not shown). Therefore, spectra were recorded at 1 °C and deconvoluted assuming the presence of two pure components (α -helix and random coil) (Fig. 15). While the CD spectra recorded at 1 °C show spectral features typical for α -helical structures with ellipticity minima at around 208 nm and 222 nm (Fig. 15, black lines), heating to 65° C results in both cases in a hypsochromic shift of the minimum to about 200 nm clearly indicative for the presence of pure random coiled structures and thus complete unfolding/denaturation of the peptides (Fig. 15, red lines). The CD predicted secondary structure content of the two peptides at 1 °C is given in Table 10. The far-UV CD-spectra used for CCA deconvolution [130] are depicted in Fig 15 (panel A and B; black line). At 1 °C both peptides are partially helical with 41 % (PpSB1-J α) and 56 % (PpSB2-J α) α -helix content.

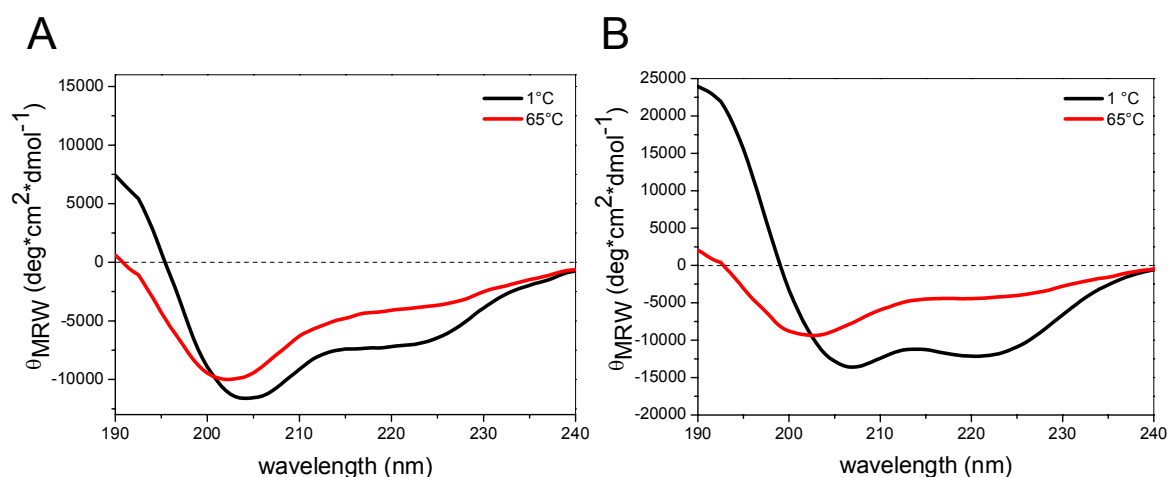


Figure 15: Far-UV CD spectra of PpSB1-J α (**A**) and PpSB2-J α (**B**) Samples were diluted in 20 mM sodium phosphate buffer (pH 7.5) supplemented with 1 mM NaCl to an final concentration of 60 μ M (1.5 mg/ml). All spectra shown in A and B were recorded at 1 °C (black line) and at 65 °C (red line).

To study thermal unfolding sequential CD spectra were collected for the two peptide samples equilibrated at 1 °C and subsequently for the same samples heated stepwise up to > 65 °C with a constant ramp rate of 1 °C / minute.

Table 10: CCA deconvolution of far-UV CD spectra recorded for the PpSB1-J α and PpSB2-J α peptides

	CD derived secondary structure content ^{&}	
	number amino acids / (%)	
	α -helix	Random coil
PpSB1-J α (23 amino acids)	9 / (41)	14 / (59)
PpSB2-J α (29 amino acids)	16 / (56)	13 / (44)

[&]: derived from spectra recorded at 1 °C

Upon heating clear loss of negative ellipticity at 222 nm and 208 nm can be detected, suggestive for the loss of α -helical secondary structure (Fig. 16). Eventually, heating results the formation of spectral species with a blue-shifted minimum, resembling the CD spectra of random coiled model peptides [150]. The presence of one isodichroic point at 203 nm is supportive of a two-state transition (Fig. 16).

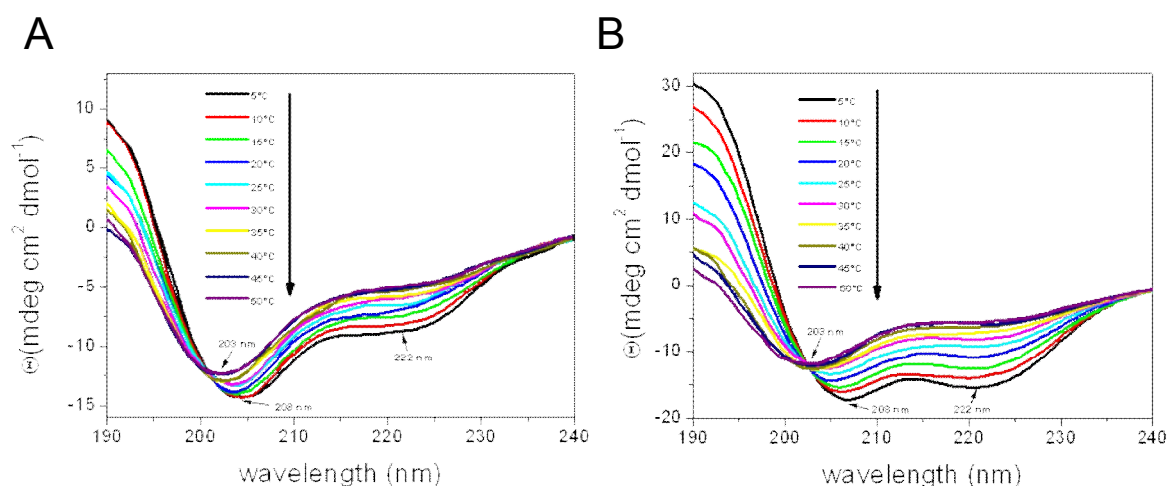


Figure 16: Far-UV CD spectra of the PpSB1-J α (A) and PpSB2-J α peptide (B) samples recorded between 5 °C - 50 °C. Samples were equilibrated at 1 °C. Subsequently CD spectra were recorded for the same sample heated stepwise to > 65 °C using a constant ramp rate of 1 °C per minute. For clarity only the spectra recorded from 5 °C - 50 °C are shown.

Interestingly, the thermal unfolding process is fully reversible for both peptides with both the unfolding and folding transition displaying the same transition behavior (Fig. 17).

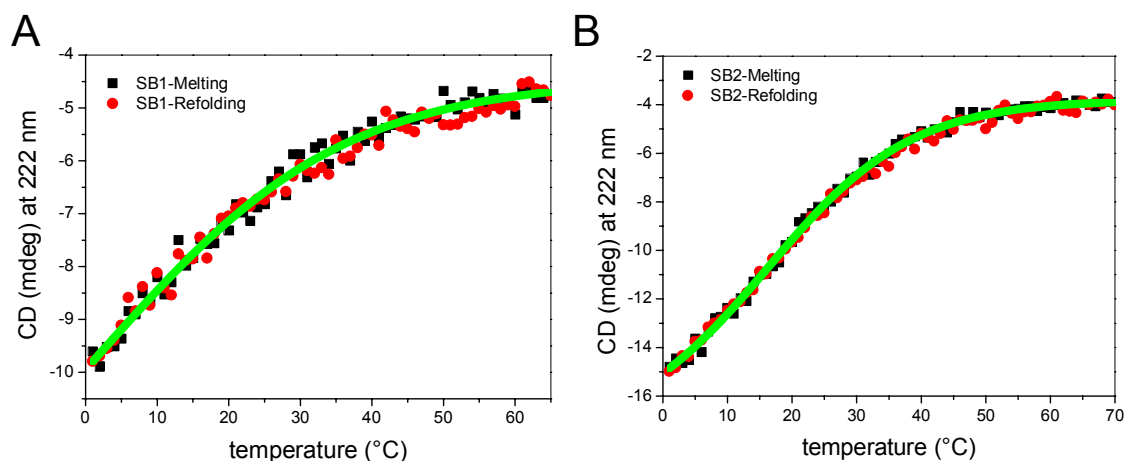


Figure 17: Unfolding/”Melting” (black) and refolding (red) curves recorded for the PpSB1-J α (A) and PpSB2-J α (B) peptides. The depicted curves were extracted from the respective far-UV CD spectra by plotting the ellipticity value at 222 nm against the incubation temperature. Samples were equilibrated at 1 °C. Subsequently CD spectra were recorded for the same sample heated stepwise to > 65 °C using a constant ramp rate of 1 °C per minute. After complete unfolding was achieved CD spectra were recorded for the same samples cooled stepwise to 1 °C using a constant ramp rate of 1° per minute.

Under the premise of complete reversibility, while assuming a two-state transition, thermodynamic parameters for the PpSB1-J α and PpSB2-J α unfolding/folding equilibrium can be extracted from CD data [157] (Table 11). Details regarding the analysis of CD melting

curves and the extraction of thermodynamic parameters can be found in the corresponding Materials and Methods section (2.9.4.1).

Table 11: Thermodynamic parameters derived from CD unfolding/refolding curves^s

Peptide	T _m (°C)	ΔH (kcal mol ⁻¹)	ΔS (kcal K ⁻¹ mol ⁻¹)	ΔG (kcal mol ⁻¹) ^{&}
PpSB1-Jα (23 amino acids)	4	9.3	0.0336	0.26
PpSB2-Jα (29 amino acids)	16	14.3	0.0494	0.87

^s: CD spectra of both peptides were determined in thermal equilibrium. Heating was performed from 1 °C to 65 °C until complete unfolding was observed

With a melting temperature (T_m) of 16 °C, PpSB2-Jα appears slightly more stable against thermal unfolding than the PpSB1-Jα peptide (T_m = 4 °C). At 1 °C, PpSB1-LOV is marginally stabilized against unfolding with a Gibbs free energy of unfolding of ΔG = 0.26 kcal·mol⁻¹ (ΔH = 9.3 kcal·mol⁻¹, ΔS = 0.0336 kcalK⁻¹mol⁻¹). For the thermally more stable peptide, PpSB2-Jα, the Gibbs free energy ΔG of unfolding is increased to 0.87 kcal mol⁻¹ (ΔH = 14.3 kcal·mol⁻¹, ΔS = 0.0494 kcalK⁻¹mol⁻¹).

CD spectroscopic studies on the isolated PpSB1-Jα and PpSB2-Jα peptides in aqueous solution reveal independent folding of the two peptides. Both peptides assume α-helical conformation which is stabilized at lower temperatures. Analysis of CD derived melting/refolding curves suggests that both peptides form thermally rather instable helical structures.

3.1.3.3 Addition of the helix-stabilizing agent 2,2,2-trifluoroethanol (TFE) is necessary to stabilize the two peptides for NMR studies

In order to infer the atomic resolution structure of the respective peptides using solution NMR techniques, ¹H-NMR spectra would have to be recorded at lower temperatures. In particular for the PpSB1-Jα peptide this would be experimentally challenging, because this peptide has apparently not reached its fully folded conformation even at 1 °C. Moreover, initial NMR studies revealed severe line broadening and at low temperatures (5 °C) resulting in poorly

resolved spectra (data not shown). In order to stabilize the helical conformation of both peptides at room temperature, 2,2,2-trifluoroethanol (TFE) an helix stabilizing agent was added to the peptide preparations. Hereby, TFE is not a helix-inducing solvent in the sense that it will induce helix formation independently of the sequence. It is rather a helix-enhancing co-solvent that stabilizes helices in regions with some α -helical propensity [151, 152]. To verify proper folding at 25 °C in the presence of minimal amounts of TFE, CD spectra were recorded for the peptide samples at 25 °C while the TFE concentration in the sample was sequentially increased (Fig. 18).

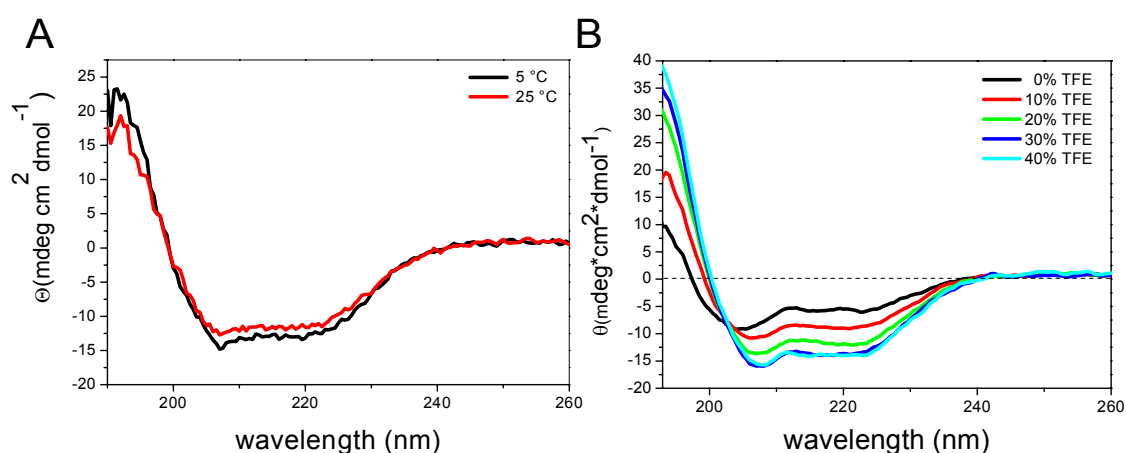


Figure 18: (A) Far-UV CD spectra of the PpSB1-J α peptide in the presence of 10 % (v/v) TFE measured at 5 °C (black line) and 25 °C (red line). (B) Far-UV CD spectra of the PpSB2-J α peptide recorded at 25 °C in the absence (0 %) and presence of different concentration of TFE.

In this way the minimal concentration of TFE needed to promote complete folding of the respective peptide was determined. While for the PpSB1-J α addition of 10 % (v/v) TFE was sufficient to induce complete folding at 25 °C (Fig. 18, panel A), 25 % (v/v) of TFE had to be added to stabilize folding of the PpSB2-J α peptide at the same temperature (Fig. 18, panel B). Above those TFE concentrations no further spectral changes were detectable for the two peptides.

TFE stabilizes the helical conformation of both peptides at higher temperatures (above the melting point of both peptides in aqueous solution). Thus by supplementing the sample with 10 % (v/v) TFE in case of the PpSB1-J α and 25 % (v/v) TFE in case of the PpSB2-J α , the peptides can be sufficiently stabilized for NMR studies to be conducted at 20 °C.

3.1.3.4 NMR experiments

Peptides samples were prepared as described in 2.11.1 and 3.1.3.1. As mentioned previously NMR spectra for both peptides at 5 °C displayed very broad peaks. Therefore, TFE was added to both peptide samples to increase helical propensity at a measuring temperature of 20 °C. The structuring properties of TFE contribute to an increase in the separation performance. For the PpSB1-J α peptide, 10 % (v/v) TFE was used and for the PpSB2-J α peptide 25 % (v/v) TFE was used (3.1.3.3). 2D ^1H - ^1H NOESY and 2D ^1H - ^1H TOCSY spectra were recorded and used for sequential resonance assignment.

Hereby, in the 2D ^1H - ^1H TOCSY spectrum the chemical shifts of pairs of protons that are within the same amino acid can be assigned, whereas the 2D ^1H - ^1H NOESY spectrum identifies the chemical shifts of pairs of protons that are close together within the peptide structure. Sequential assignments were obtained by analyzing sequential $\text{H}^{\text{N}}_{(i)}\text{-H}^{\text{N}}_{(i+1)}$, $\text{H}^{\alpha}_{(i)}\text{-H}^{\alpha}_{(i+1)}$ and $\text{H}^{\beta}_{(i)}\text{-H}^{\text{N}}_{(i+1)}$ NOE correlations [138]. NMR chemical shift information for PpSB1-J α and PpSB2-J α has been deposited at the Biological Magnetic Resonance Databank (BMRB) and the accession numbers: 18159 and 18160 and is listed in Table A.3 and Table A.4 in the appendix.

3.1.3.5 Structure determination of PpSB1/PpSB2-J α peptide by NMR

Peptide structures were calculated as described in 2.11.8, by using total 537 restraints for PpSB1-J α , of which 485 were unambiguous and 749 restraints for PpSB2-J α , with 721 restrains being unambiguous.

Quality, structural, restraint and violation statistics for the complete ensemble of 10 structures for the two peptides are presented in Table 12.

The final ensemble of the 10 structure (Fig. 19) was selected by energy from an ensemble of 250 structures. The RMSD of the backbone atoms from the mean structure was 0.053 Å for PpSB1-J α (ordered region Arg121-Ala134) and 0.191 Å for PpSB2-J α (ordered region Ala121-Gln142). No NOE violations were observed. Fig. 19 depict the overlay of the 10 ensemble structures (C α -trace as ribbon drawing) superimposed over the respective helical portion (as transparent cartoon).

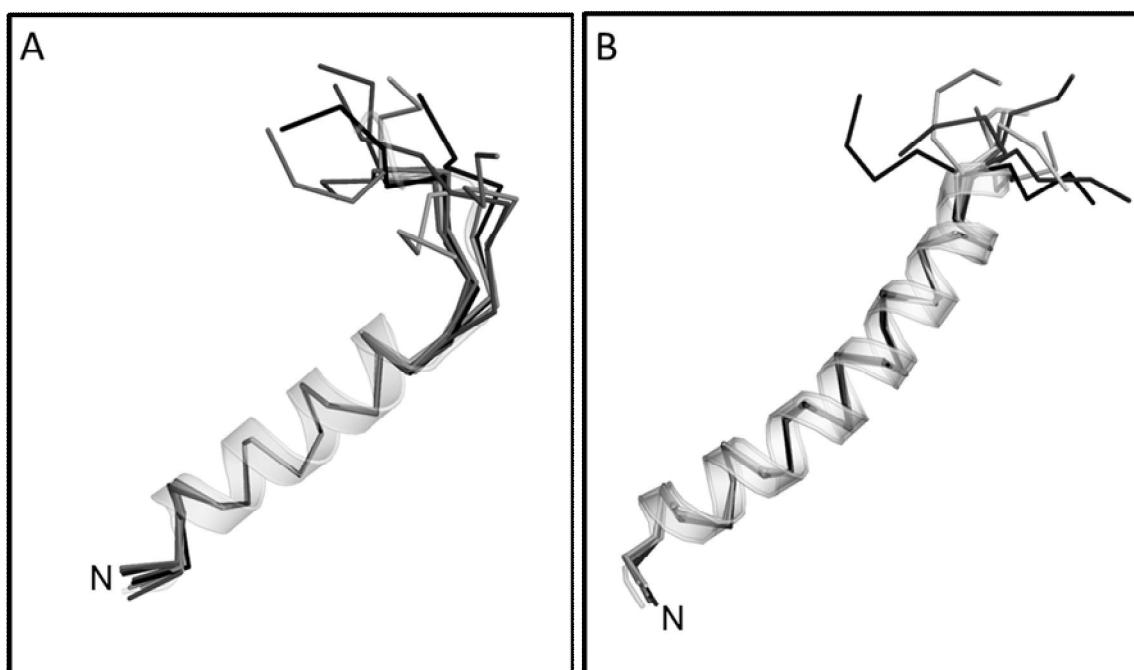


Figure 19: Ensemble of the 10 NMR derived structures of PpSB1-J α (A) and PpSB2-J α (B) selected based on energy from an ensemble of 250 structures. The polypeptide backbone of the ten lowest energy structures is shown in ribbon representation (grey to black). The energy minimized average structure is superimposed onto the ensemble (shown in white, transparent cartoon representation).

Structure calculations revealed a well-defined helical structure for both peptides (Fig. 19). Please note that NMR structural models derived from NOE-based structure calculations inevitably overestimate the real secondary structure content of the ensemble. In order to quantify α -helical secondary structure content directly from NMR chemical shift data the secondary structure propensity score (SSP) developed by Marsh and co-workers [153] was used. A SSP score at a given residue of 1 or -1 are indicative of fully formed α - or β -structure, respectively. A score of 0.5 indicates that 50 % of the conformers in the disordered state ensemble are helical at that position [153]. Fig. 20 depict the residue-wise SSP score for PpSB2-J α (panel A) and PpSB1-J α (panel B). For both peptides most residue SSP scores are positive indicative of a certain degree of α -helical propensity. Exceptions include R135 (PpSB1-J α) and K147 (PpSB2-J α), which show negative values. Average SSP scores for the ordered region of PpSB1-J α (residues 121-134) and PpSB2-J α (residues 121-142) are 0.48 and 0.44 respectively. This suggests that on average 48 % and 44 % of the conformers in the respective regions are α -helical in the disordered state ensemble.

Table 12: Structural, distance restraint and violation statistics for the 10 lowest energy structures of PpSB1-J α and PpSB2-J α

	PpSB1-Jα (119-142)	PpSB2-Jα (119-148)
<i>Experimental restraints</i>		
Total NOE distance restraints	537	749
Unambiguous	485	721
Intraresidue	190	321
Interresidue	295	400
Sequential	159	195
Short range (2-3)	108	157
Medium range (4-5)	28	48
Long range (>5)	0	0
Ambiguous	52	28
<i>Number of NOE upper bound violations</i>		
Violations > 0.5 Å	0	0
Violations > 0.3 Å	0	0
Violations > 0.1 Å	12.1 \pm 1.7	19.4 \pm 2.0
<i>Deviations from idealized covalent geometry</i>		
RMSD		
Bonds (Å)	0.004 \pm 0.0001	0.004 \pm 0.0001
Angles (deg)	0.519 \pm 0.0157	0.583 \pm 0.0146
Improper (deg)	0.965 \pm 0.1498	1.247 \pm 0.1199
<i>Deviations from experimental restraints</i>		
RMSD		
NOE restraints (Å)	0.028 \pm 0.002	0.027 \pm 0.000
<i>Ramachandran plot (Ensemble)</i>		
Favored region	95.9 %	95.4 %
Allowed region	100 %	100 %
Disallowed region	0 %	0 %
<i>Ensemble RMSD (Å)</i>		
Core residues	121-134	121-142
Backbone	0.053	0.191
Heavy atoms	1.003	0.789

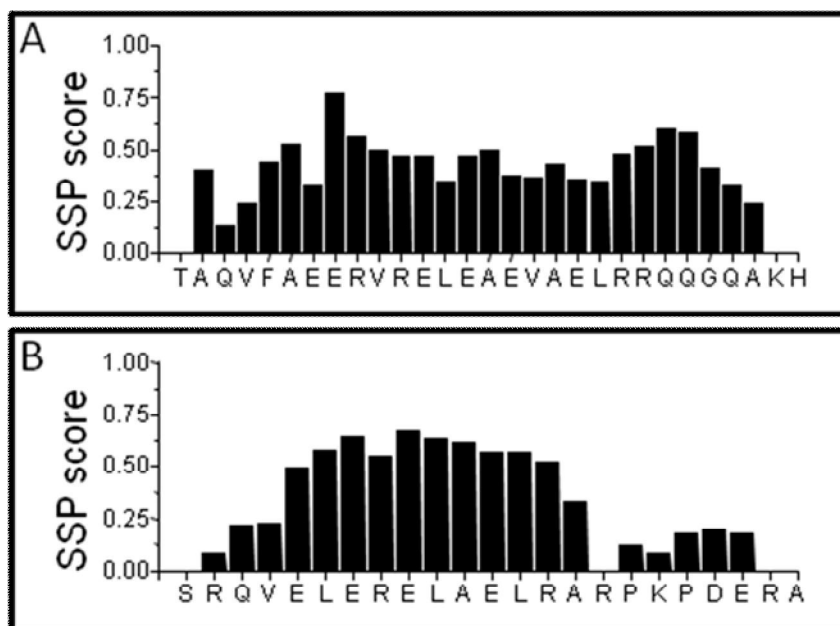


Figure 20: Secondary structure propensity scores (SSP) calculated for PpSB2-J α (A) and PpSB1-J α (B) according to Marsh *et. Al* [153] using the respective chemical shift information. A SSP score of 1 is indicative of fully formed α -helical structure. A score of 0.5 indicates that 50% of the conformers in the disordered state ensemble are helical at a given residue position.

While SSP scores provide a more realistic measure for the true α -helical propensity of the sample, they unfortunately do not allow a clear cut estimation of e.g. the percentage of residues in an α -helical conformation. Given the above discussed average SSP score of 0.48 and 0.44 for the ordered region of both peptides we can set an arbitrary minimum threshold for the propensity of defining a residue in α -helical conformation of about 0.4. Assuming this arbitrary a SSP threshold, 9 residues out of 22 for PpSB1-J α (41 %) and 16 out of 28 residues for PpSB2-J α (57 %) possess a strong tendency for α -helical secondary structure. This translates roughly to 3 and 5 helical turns in structure for PpSB1-J α and PpSB2-J α , respectively.

NMR studies provide atomic resolution structural models for both isolated PpSB1-J α and PpSB2-J α peptides. As already suggested by CD spectroscopy, both peptides adopt an α -helical conformation, whereas the helical structure of PpSB2-J α is extended by roughly two helical turns compared to PpSB1-J α .

3.2 Solution NMR experiments using the full-length PpSB1-LOV protein

3.2.1 Optimization of expression and purification of PpSB1-LOV for NMR spectroscopic studies

The most important prerequisite for NMR structural studies is the availability of large amounts of uniformly labeled structurally homogenous target protein. As isotopically labeled protein can only be produced using minimal media containing the respective labeled N- and C- sources, expression conditions using those media have to be optimized for each target protein. Different strategies for the production of ^{15}N - and ^{13}C labeled PpSB1-LOV protein will be described in the following.

3.2.1.1 Optimization of expression and purification conditions – Expression in *E. coli* BL21 (DE3)

Previous experiments had suggested autoinduction full-media (AI-FM, see 2.5.2.2) as optimally suited to obtain high amounts of fully FMN loaded PpSB1-LOV protein from *E. coli* BL21 (DE3) expression cultures [49]. Therefore, initial experiments were carried out using a minimal media substitute for AI-FM. Using such a media, isotopic labeling is possible by substituting glucose and glycerol by ^{13}C -glucose and ^{13}C -glycerol as carbon sources, and by substituting ammonium chloride for the respective ^{15}N derivative [117]. The expression of PpSB1-LOV was thus carried out initially in *E. coli* BL21 (DE3) over different time periods using autoinduction minimal media (AI-MM, see 2.9.1.5.4) to optimize the overall protein yield. The corresponding SDS gel is shown in Fig. 21. The optimal time for expression was found to be 72 hours (marked by an asterisk in Fig. 21).

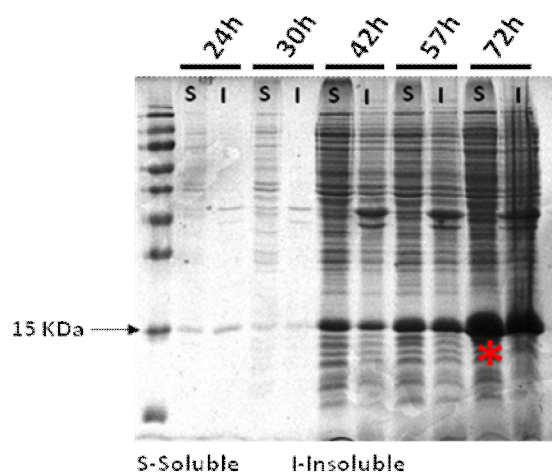


Figure 21: Coomassie stained SDS-PAGE documenting the expression of PpSB1-LOV in *E. coli* BL21 (DE3) using autoinduction minimal media over a time period of 72 hours. The arrow indicates the PpSB1-LOV band. Lane 1: Molecular weight standard. Lane 2: expression after 24 h. Lane 3: expression after 30 h. Lane 4: expression after 42 h. Lane 5: expression after 57 h. Lane 6: expression after 72 h.

In order to analyze the quality of the expressed LOV protein, His-tagged PpSB1-LOV was purified from *E. coli* BL21 (DE3) cells as described in the materials and methods section (2.9.3.1.2). The IMAC purified PpSB1-LOV protein was dialyzed (2.9.3.2) against 10 mM sodium phosphate buffer (pH 8.0) supplemented with 10 mM NaCl. Subsequently, the chromophore-load and chromophore-composition was analyzed using UV/Vis spectrophotometry (2.9.4.2) and HPLC (2.9.3.9).

Using the aforementioned expression and purification protocol 15 mg of pure protein per litre *E. coli* expression culture could be obtained.

3.2.1.1.1 Chromophore-load of the *E. coli* BL21 prepared PpSB1-LOV protein analyzed by UV/Vis spectrophotometry

In order to verify complete loading of the PpSB1-LOV protein with its light sensitive flavin chromophore UV/Vis spectra (200-650 nm) were recorded for the purified protein in its dark-state. The respective protein sample was diluted in 10 mM sodium phosphate buffer (pH 8.0) supplemented with 10 mM NaCl to a final absorbance of about 0.2 at 450 nm (Fig. 22). The same buffer was used as a reference.

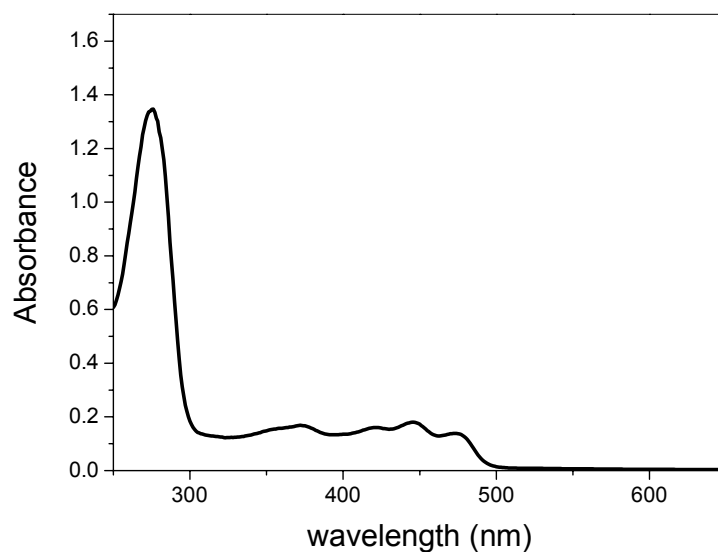


Figure 22: UV-Vis absorption spectrum of purified PpSB1-LOV in its dark adapted state. The PpSB1-LOV protein was expressed in *E. coli* BL21 (DE3) and purified as described in 2.9.3.1.2.

From the flavin-specific absorbance of the sample at 450 nm in relation to the protein specific absorbance determined at 280 nm the chromophore-loading of PpSB1-LOV can be estimated (see Materials and Methods for details, 2.9.4.2.2). The purified protein obtained from *E. coli* BL21 (DE3) expression cultures was only loaded to approx. 45 % with the light sensitive flavin chromophore.

3.2.1.1.2 Chromophore-composition of the *E. coli* BL21 prepared PpSB1-LOV protein analyzed by HPLC

Depending on the expression conditions, LOV proteins can accept the three naturally occurring flavin-type chromophores (riboflavin, FAD and FMN) to a variable degree [49]. In order to evaluate the flavin chromophore-composition of the PpSB1-LOV protein obtained from *E. coli* BL21 (DE3) expression cultures, a HPLC analysis was performed (2.9.3.9). The protein-bound chromophore was released by heat denaturation (2.9.3.9). Separation of Riboflavin, FAD and FMN was achieved by reverse-phase HPLC (2.9.3.9). FMN, FAD and riboflavin were used as reference compounds. A typical HPLC chromatogram is shown in Fig. 23. Peak retention times were as follows: FAD: 9.7 min, FMN: 10.6 min and Riboflavin: 12.7 min.

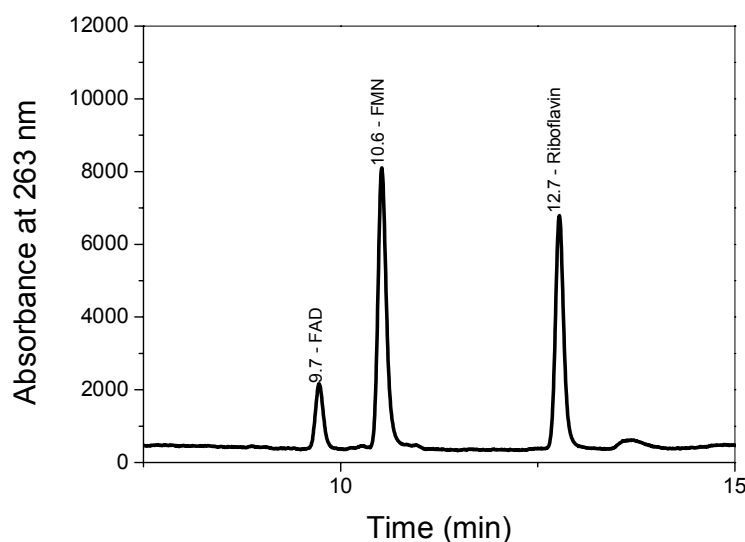


Figure 23: HPLC analysis of the flavin pool released by heat denaturation from the PpSB1-LOV sample. The PpSB1-LOV protein was expressed in *E. coli* BL21 (DE3) and purified as described in 2.9.3.1.2.

HPLC analysis (Fig. 23) revealed that the PpSB1-LOV preparation obtained from *E. coli* BL21 (DE3) expression cultures contained a mixture of all three natural flavins, i.e. FAD: ~9.8 %, FMN: ~49.5 % and riboflavin: ~40.7 %.

Chromophore-load and -composition analyses suggested that the protein preparation obtained from *E. coli* BL21 (DE3) is not homogenous. In particular, significant amounts of apo-protein is present in the preparation and a mixture of structurally different flavin species is loaded on the protein. This might result in structural inhomogenities compromising NMR structural studies.

3.2.1.2 Optimization of expression and purification conditions - Expression in *E. coli* CmpX131 (DE3)

In order to increase chromophore loading and improve the incorporation of the native FMN chromophore into the heterologously produced protein, a recently described riboflavin auxotrophic strain *E. coli* CmpX131 (DE3) [104] was used. In the CmpX131 (DE3) strain,

exogenously provided riboflavin cofactors are incorporated into the overproduced protein during protein biosynthesis. The strain carries a deletion in the *ribC* gene, responsible for riboflavin biosynthesis [104]. Moreover, the strain contains a the riboflavin transporter gene *ribM* (*pnuX*) from *Corynebacterium glutamicum* [154], which allows efficient import of riboflavin from the culture medium (Fig. 26, panel A) [104]. As an *E. coli* C41 (DE3) derivative the strain facilitates P_{T7} promoter dependent gene expression thus enabling the use of conventional expression vectors such as the pET vector system for the overproduction of flavoproteins [104]. Additionally the strain enables the import of non-natural flavoprotein cofactors such as e.g. roseoflavin, which are then incorporated into the target flavoprotein of choice [104]. Since the strain lacks the capability to biosynthetically produce riboflavin, no growth will be observed when the strain is in grown in media lacking an external riboflavin source (Fig. 24, panel B). In turn, by supplementing the media with sufficient amounts of riboflavin, the strain will never encounter flavin limitation under PpSB1-LOV overexpression conditions. This might result in improved chromophore-loading and flavin content.

A PpSB1-LOV protein sample was prepared from *E. coli* CmpX131 HD-MM expression cultures (2.9.1.5.3) and purified as described in 2.9.3.1. Similar experiments were conducted using autoinduction minimal media (AI-MM) (2.9.1.5.4), however, we eventually shifted to using high cell density IPTG induction medium (HD-MM) as it yielded more protein and is less expensive, because no ^{13}C -glycerol is needed for growth. Please note that 500 mg of ^{13}C labelled glycerol cost about 300 Euro. In order to produce one litre AI-MM 5 g of ^{13}C -glycerol are needed.

With 30 mg of pure protein per litre of *E. coli* expression culture the recombinant protein yield was twice as high compared to the autoinduction (AI-MM) expression system (3.2.1.1).

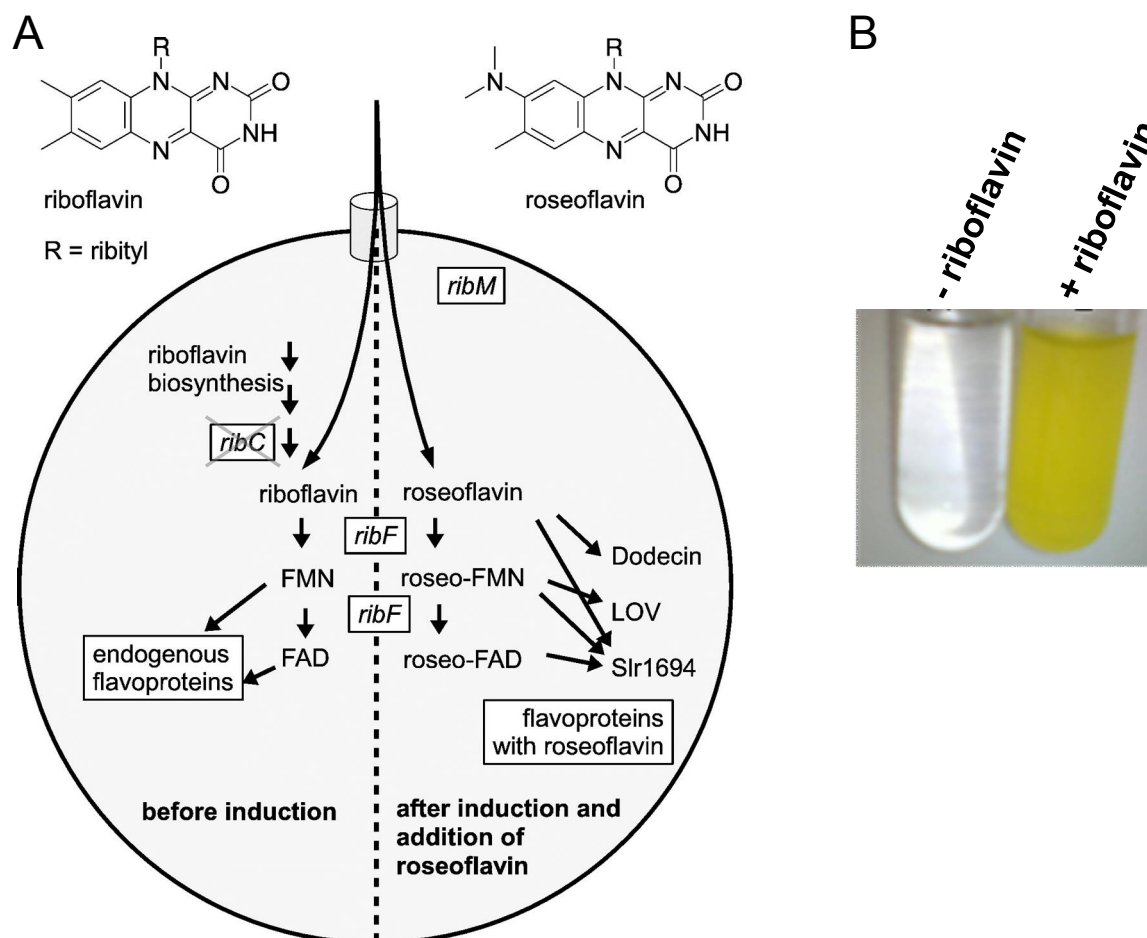


Figure 24: (A) Schematic representation of the *E. coli* CmpX131 (DE3) strain illustrating the concept for the *in vivo* generation of flavoproteins containing flavin analogs such as roseoflavin [104]. (B) *E. coli* CmpX131 (DE3) grown in autoinduction minimal medium (AI-MM, see 2.9.1.5.4) with (+) and without (-) riboflavin (50 μ M). In media lacking riboflavin no growth is observed.

The drawback of this expression system is that the FMN chromophore of the overproduced LOV protein will remain unlabeled even when ^{13}C - and ^{15}N - labeled nitrogen and carbon sources are provided in the minimal media. To achieve loading of the overproduced LOV protein with isotopically labeled FMN, respectively ^{13}C - and ^{15}N - labeled riboflavin would have to be supplied to the medium. However, uniformly labeled riboflavin is not commercially available.

3.2.1.2.1 Chromophore-load and -composition of the from *E. coli* CmpX131 prepared PpSB1-LOV protein

To estimate chromophore-loading the same procedure as above (3.2.1.1.1) also described in 2.9.4.2 (Materials and Methods) was used. The corresponding UV/Vis absorbance spectrum is

shown in Fig. 25 (panel A). Evaluation of the respective UV/Vis spectrum suggests complete loading of the *E. coli* CmpX131 (DE3) derived PpSB1-LOV preparation with the light sensitive flavin chromophore. The flavin composition of the respective preparation was evaluated as above (3.2.1.1.2) also described in section in 2.9.3.9 [45]. The corresponding HPLC chromatogram (Fig. 25, panel B) revealed that the PpSB1-LOV preparation obtained from *E. coli* CmpX131 (DE3) binds predominately FMN (96.7 %) and contains only trace amounts of FAD (2.1 %) and riboflavin (1.2 %).

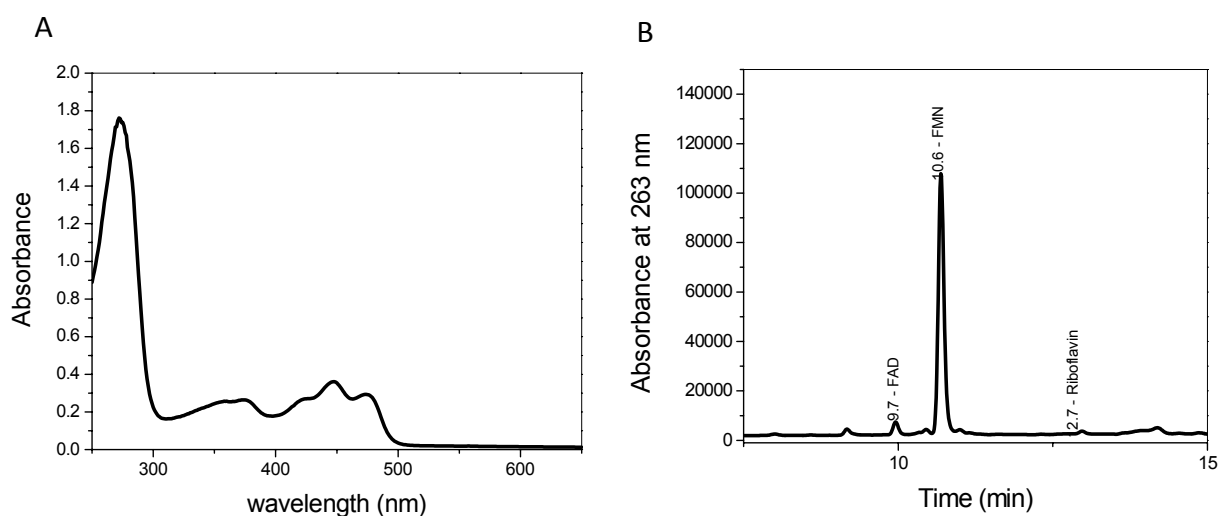


Figure 25: (A) UV/Vis absorbance spectrum of the dark-adapted state of PpSB1-LOV purified from HD-MM (2.9.1.5.3) grown *E. coli* CmpX131 (DE3). (B) HPLC analysis of the flavin pool released by heat-denaturation from the PpSB1-LOV sample. The PpSB1-LOV protein was expressed in *E. coli* CmpX131 (DE3) using HD-MM (2.9.1.5.3) media and purified as described in 2.9.3.1.2.

The use of the *E. coli* CmpX131 (DE3) expression strain in combination with a high-cell density IPTG induction overexpression system improves both the yield and homogeneity of the PpSB1-LOV preparation to a degree suitable for the production of high amounts of isotopically labeled PpSB1-LOV protein for NMR studies.

3.2.2 Initial sample preparation for NMR studies

Uniform ^{15}N - or ^{15}N - and ^{13}C - labeling was achieved by using the *E. coli* CmpX131 (DE3) expression strain grown in HD-MM (2.5.3.1) [117]. The recombinant protein was purified as described in the previous chapter (3.2.1.2).

After cell lysis (2.9.3.1.1) the His-tagged PpSB1-LOV protein was purified from the soluble extract by IMAC on Ni-NTA Superflow resin (QIAGEN, Hilden, Germany) as described previously (see Materials and Methods 2.9.3.1.2). The purified protein was dialyzed overnight against 10 mM sodium phosphate buffer (pH 8.0) supplemented with 10 mM NaCl and 1 mM DTT (2.6.3.3). The N-terminal His-tag was cleaved by proteolytic digestion using biotinylated Thrombin (2.9.3.3). The untagged PpSB1-LOV protein was obtained by passing the Thrombin digested sample through Ni-NTA column to remove the cleaved tag. Biotinylated Thrombin was removed using Streptavidin agarose (2.9.3.3). Flow-through fractions containing PpSB1-LOV without tag were pooled and dialyzed (2.9.3.2) against 10 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl and 1 mM DTT. Dialyzed protein was concentrated to 0.2-0.5 mM by spin filtration using a Vivaspin (10 kDa MWCO) centrifugal concentrator (Sartorius AG, Goettingen, Germany) (2.9.3.4). Fig. 26 depicts the corresponding SDS gel illustrating successful His-tag removal by proteolytic digestion.

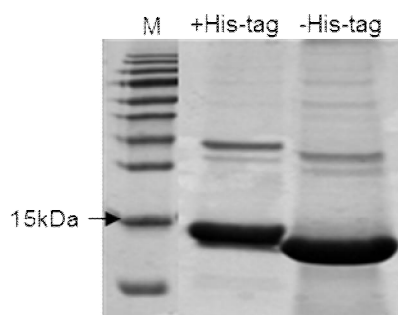


Figure 26: SDS-PAGE analysis of the proteolytic cleavage of the PpSB1-LOV His-tag by Thrombin protease. Lane 1: molecular weight standard. Lane 2: PpSB1-LOV before cleavage of the His-tag. Lane 3: PpSB1-LOV protein after removal of the tag and purification.

3.2.2.1 Blue-light sensitivity of the isotopically labeled PpSB1-LOV sample

In order to verify photosensitivity of the purified isotopically labeled PpSB1-LOV protein sample UV/Vis spectra were recorded for the dark-state of the protein (Fig. 27, black line). Subsequently the sample was illuminated for 30 s using a blue-light emitting LED lenser (450 nm) V8 lamp (Zweibrüder- Optoelectronics GmbH, Solingen, Germany). All measurements were performed at 20 °C. The protein sample contained 10 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl and 1mM DTT. Upon illumination with blue light, the flavin absorption band in the visible region at around 450 nm

decreases giving rise to a new maximum at around 375 nm (Fig. 27, red line). Those spectral changes are characteristic for the formation of a covalent flavin-C4a cysteinyl thiol adduct between protein and chromophore (see 1.2.4).

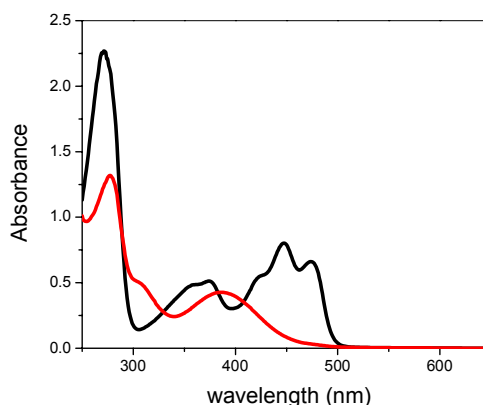


Figure 27: UV/Vis absorbance spectra of the dark (black) and light adapted state (red) of PpSB1-LOV purified from HD-MM (2.9.1.5.3) grown *E. coli* CmpX131 (DE3).

3.2.2.2 Matrix assisted laser desorption/ionization mass spectrometry (MALDI-TOF MS)

To verify complete ^{15}N - or ^{15}N - ^{13}C labeling as well as tag-cleavage additionally MALDI-TOF MS [155] was used (see Materials and Methods 2.9.3.10 for details). All measurements were performed at the MPI for Bioinorganic Chemistry (Mülheim a.d. Ruhr) in cooperation with Wolfgang Gärtner and Norbert Dickmann. For MALDI-TOF MS experiments, the isotopically-labeled protein samples were dialyzed against 10 mM Tris-HCl (pH 6.3) containing 10 mM NaCl. Fig. 28 depicts the MALDI-TOF mass-spectra of the corresponding PpSB1-LOV protein samples. The spectrum obtained for the ^{15}N -labeled protein sample shows two sharp peaks corresponding to the single (1z) and double (2z) ionized PpSB1-LOV molecule, with masses of 8455.2 Da (2z) and 16927.6 Da (1z) (Fig. 28, panel A). The average mass of the PpSB1-LOV molecule derived from the mean of the two mass peaks is thus 16919 Dalton. This value exactly matches the theoretical isotopically averaged molecular weight (16919.1 Da) for the fully ^{15}N -labeled protein estimated from the corresponding PpSB1-LOV sequence by using the protein mass calculator available at the Scripps research institute (<http://www.scripps.edu/cgi-bin/cdputnam/protcalc3>)

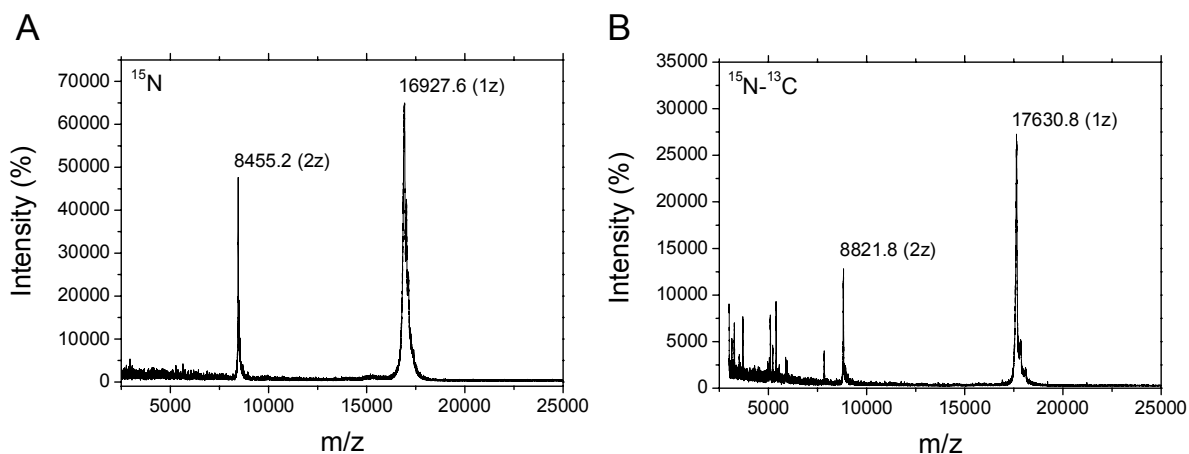


Figure 28: MALDI-TOF mass-spectra of uniformly ^{15}N - (A) and ^{15}N - ^{13}C labeled (B) PpSB1-LOV protein samples. Before MALDI-TOF MS experiments the His-tag was removed by proteolytic cleavage with Thrombin and the purified sample was dialyzed against 10 mM Tris-HCl (pH 6.3).

The spectrum obtained for the ^{15}N - ^{13}C double-labeled protein sample shows two sharp peaks corresponding to the single (1z) and double (2z) ionized PpSB1-LOV molecule, with masses of 8821.8 Da (2z) and 17630.8 Da (1z). The average mass of the PpSB1-LOV molecule derived from the mean of the two mass peaks is thus 17637.2 Dalton. This value matches the theoretical isotopically averaged molecular weight (17638.7 Da) for the fully ^{15}N - ^{13}C labeled protein estimated from sequence as described above.

Using the above described expression and purification strategy (3.2.1.2) suitable amounts of uniformly ^{15}N - or ^{15}N - ^{13}C labeled PpSB1-LOV protein can be obtained. Proteolytic cleavage of the His-tag is possible and the purified samples show identical light-dependent photochemistry compared to the unlabeled His-tagged protein used in previous studies.

3.2.2.3 Optimization of buffer/measuring conditions for NMR experiments

Uniformly ^{15}N - labeled PpSB1-LOV protein was produced and purified as described above (3.2.2). Initially ^1H - ^{15}N HSQC spectra were recorded for a dark adapted sample at 25 °C (Fig. 29, panel A). The ^1H - ^{15}N -HSQC spectrum was not resolved so the temperature and buffer conditions needed to be optimized. To determine the optimal measuring temperature, HSQC

spectra were recorded by varying the temperature from 25 °C to 40 °C (Fig. 29, panel A to D). Better resolved spectra could be obtained at higher measuring temperatures (i.e. 40 °C).

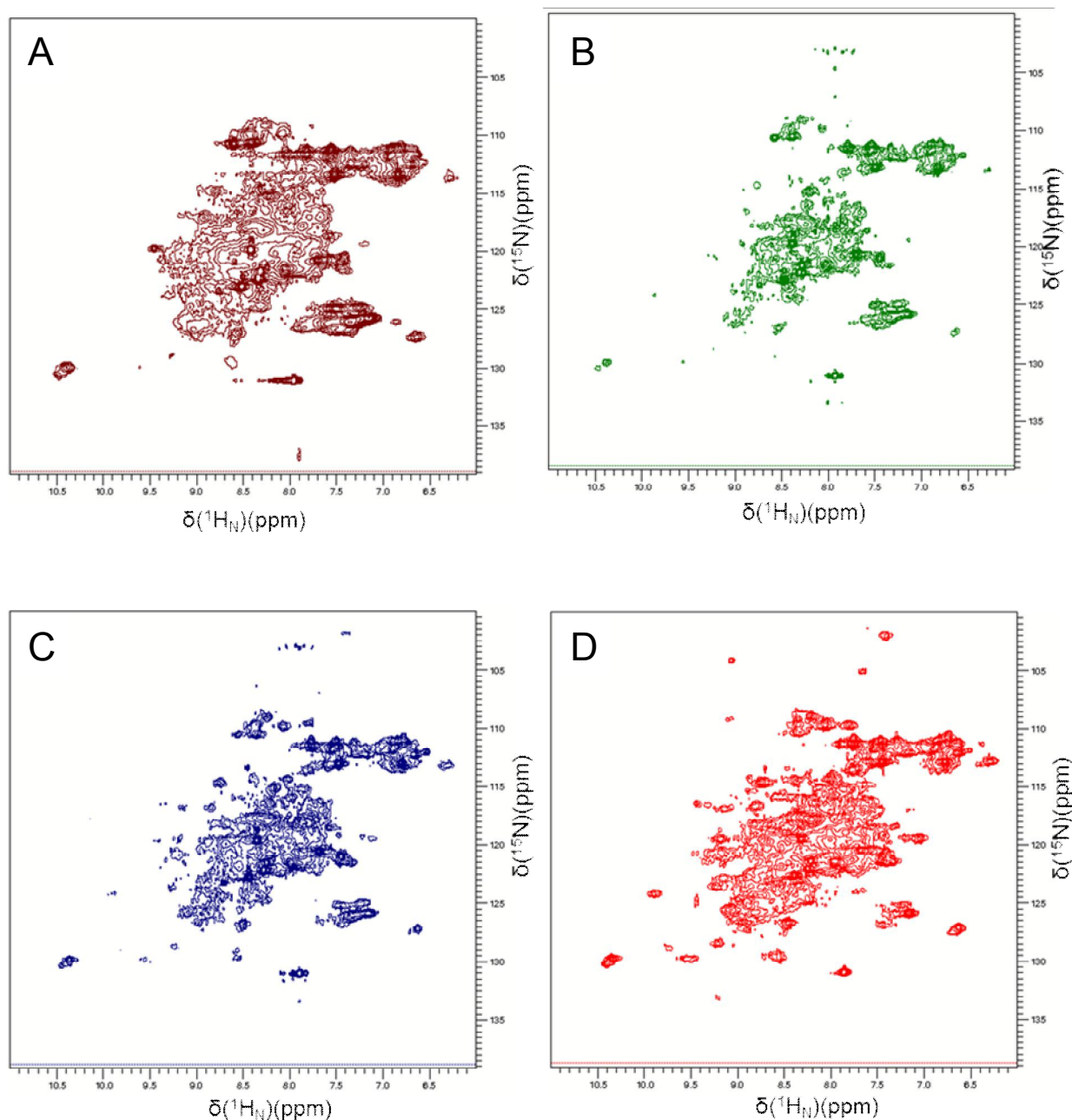


Figure 29: Measuring temperature optimization for the PpSB1-LOV protein. ^1H - ^{15}N HSQC spectra were recorded at 25 °C (A), 30 °C (B), 35 °C (C) and 40 °C (D). In all experiments NMR buffer (2.6.3.3) was containing 0.05 % (v/v) sodium azide and 10 % (v/v) D_2O . A final protein concentration of 0.5 mM was used in all experiments.

After optimization of the measuring temperature, the effect of the addition of divalent cations and chelating agents such as EDTA was tested. Therefore, EDTA, MgSO_4 and CaCl_2 were added to the respective protein sample in a final concentration of 1 mM. While addition of

MgSO₄ did not have a dramatic effect on spectra quality (data not shown) addition of 1 mM CaCl₂ to the protein sample further improved spectral resolution to a degree suitable for protein backbone resonance assignment (Fig. 30, panel B).

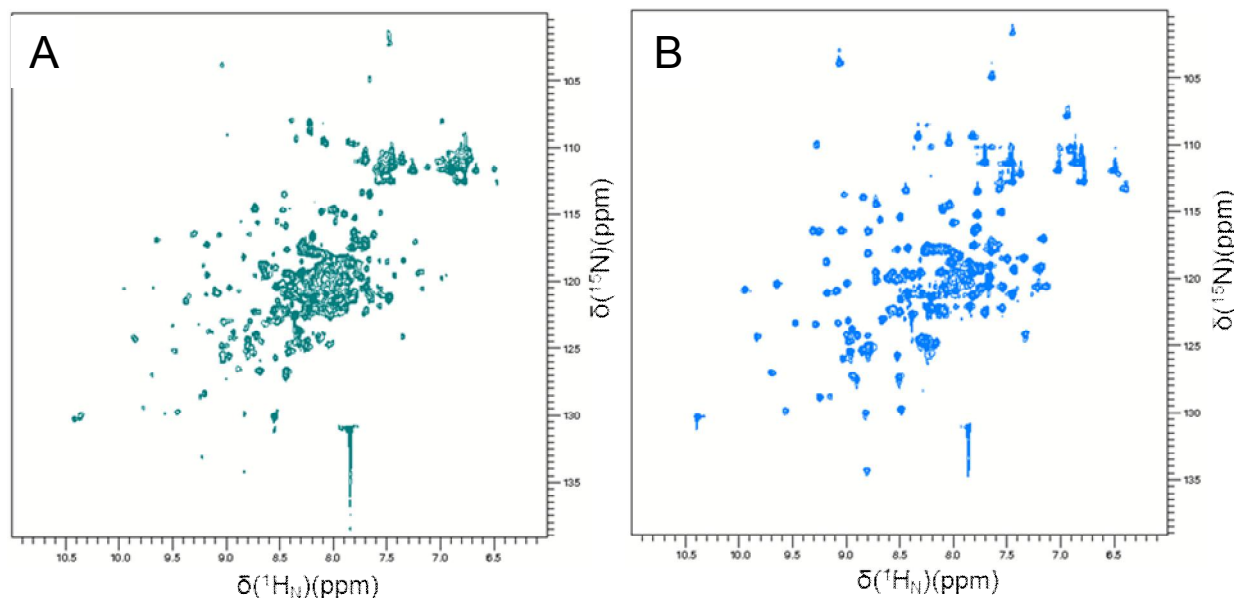


Figure 30: Buffer composition optimization for PpSB1-LOV NMR measurements. The NMR buffer (2.6.3.3) containing 0.05 % (v/v) sodium azide and as 10 % (v/v) D₂O was additionally supplemented with 1 mM EDTA (A) or with 1 mM CaCl₂ (B). All measurements were carried out at 40 °C.

3.2.2.4 Final sample preparation procedure used for all NMR experiments

Given the above observations (3.2.2.3) the sample preparation procedure was adapted in the following way: the IMAC purified (2.9.3.1.2), Thrombin digested (2.9.3.3) concentrated protein sample was supplemented with 1 mM CaCl₂ and incubated at 40 °C for 4 hours in the dark. Precipitated protein was removed by centrifugation (12000 rpm, 4 °C, 20 min) and the buffer was changed by dialysis to NMR buffer (2.6.3.3) The final sample preparation protocol is summarized schematically in Fig. 31.

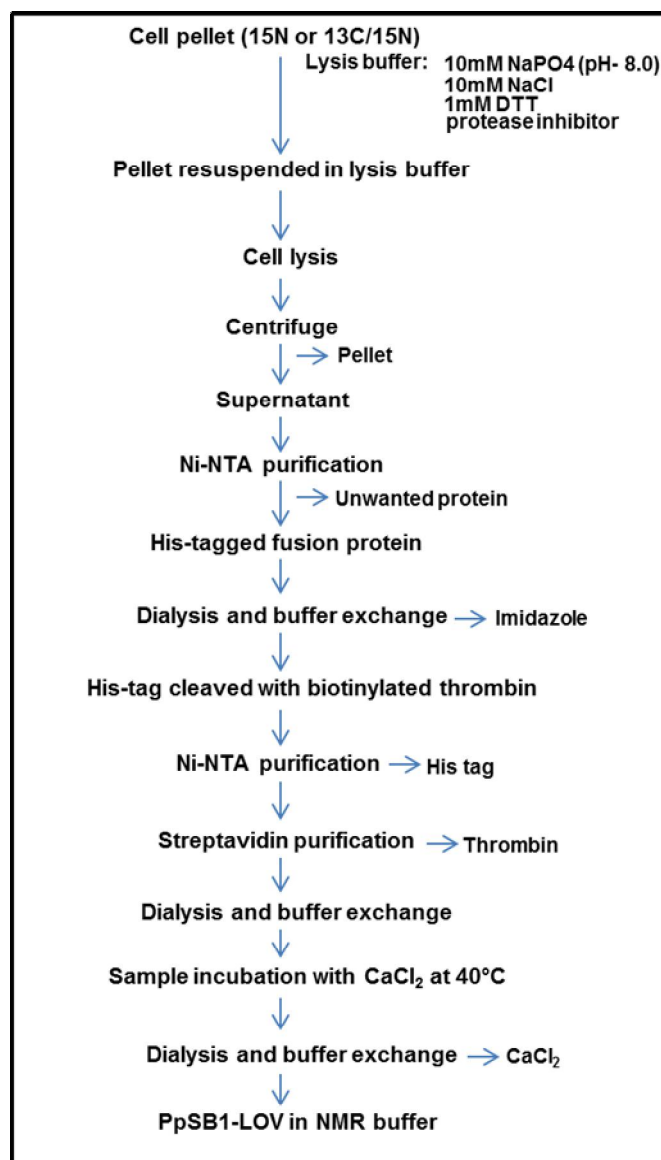


Figure 31: Schematic representation of the sample preparation procedure used for all NMR experiments.

The sample preparation, buffer composition and measuring conditions for PpSB1-LOV could be optimized to obtain well resolved ^1H - ^{15}N HSQC spectra. To obtain well resolved spectra, the sample had to be incubated for up to 4 hours at 40 °C in the presence of 1 mM CaCl_2 . Removal of the cation after heat incubation did not compromise spectral quality. Optimum resolution was obtained when spectra are recorded at 40 °C.

3.2.2.5 Final dark-state ^1H - ^{15}N HSQC spectra

Employing the above (3.2.2.4) sample preparation protocol well-resolved ^1H - ^{15}N HSQC spectra were recorded for the PpSB1-LOV protein in its dark-state (Fig. 32).

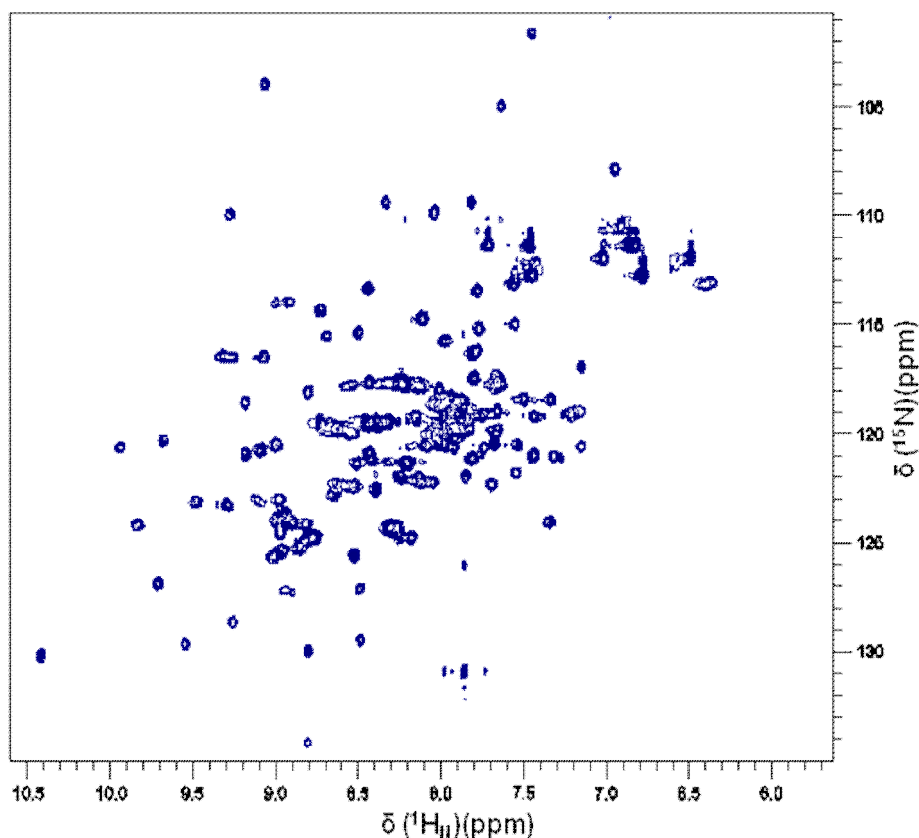


Figure 32: ^1H - ^{15}N HSQC spectra of PpSB1-LOV in the dark-state recorded using a 900 MHz NMR spectrometer at a measuring temperature of 40 °C. Sample concentration was set to 0.5 mM using 10 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl and 1 mM DTT. Additionally, the sample contained 0.05 % (v/v) sodium azide as well as 10 % (v/v) D_2O .

The spectrum is well resolved with 130, out of 138 theoretically possible backbone amide resonances, being detectable.

3.2.2.6 A custom made illumination device allows sample illumination inside the NMR spectrometer

For collection of light-state NMR spectra, a blue-light emitting LED ($\lambda_{\text{max}} = 470$ nm) was coupled via a collimator lens into a glass fibre. Using this fibre optics illumination device, the PpSB1-LOV protein sample could be directly illuminated in the shigemi tube inside the NMR

spectrometer (Fig. 33, panel A). Using a uniformly ^{15}N - labeled protein sample and the illuminating device ^1H - ^{15}N - HSQC spectra were measured for the light state of the PpSB1-LOV protein (Fig. 33, panel B).

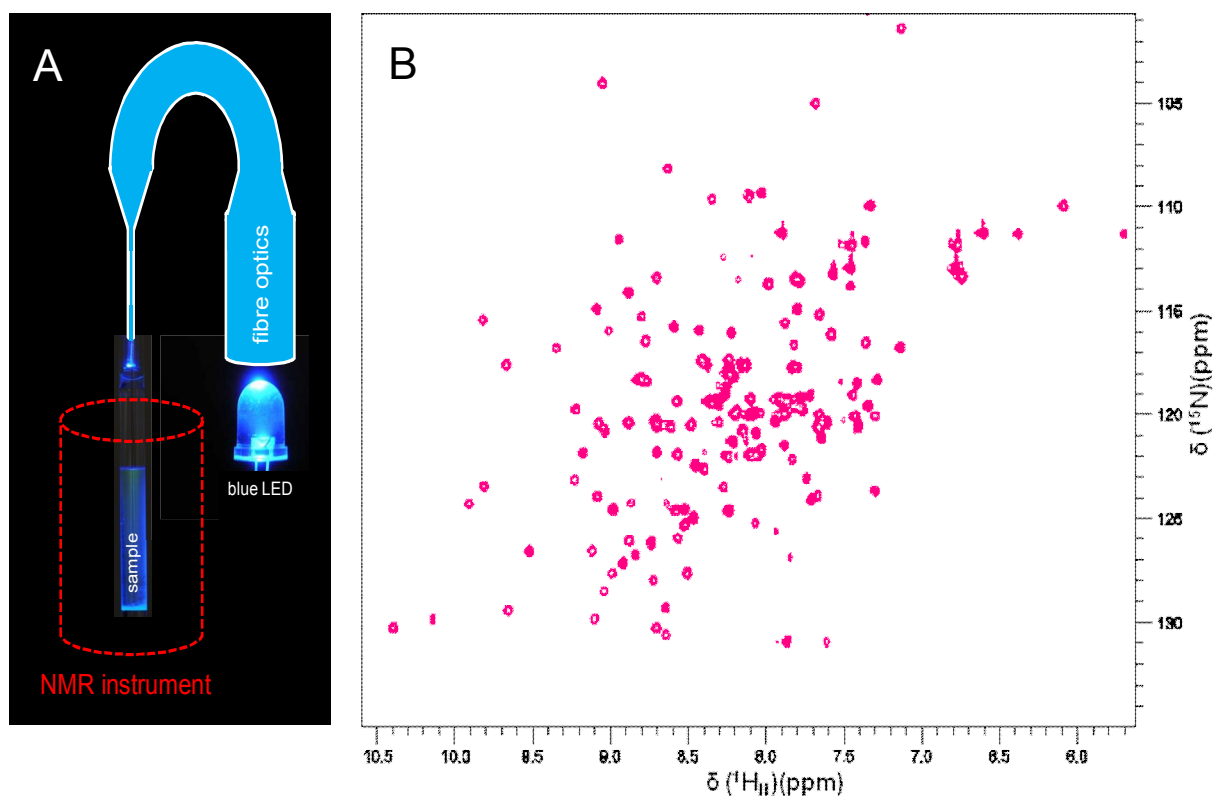


Figure 33: (A) Experimental setup for all NMR measurements of the PpSB1-LOV light-state. (B) ^1H - ^{15}N HSQC spectra of PpSB1-LOV in the light-state recorded at 900 MHz. Sample concentration was set to 0.5 mM using 10 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl and 1 mM DTT. Additionally, the sample contained 0.05 % (v/v) sodium azide as well as 10 % (v/v) D_2O . Measuring temperature: 40 °C.

In comparison to the dark-state spectrum shown in Fig. 32, light-state spectral peaks are sharper and better resolved.

3.2.2.7 Comparison of dark- and light-state PpSB1-LOV HSQC spectra suggest widespread structural changes upon illumination of the protein

As pointed out in the previous chapters well resolved HSQC spectra could be obtained for the PpSB1-LOV protein in both states. The high degree of dispersion in both the ^{15}N and ^1H dimensions is indicative of a well-folded protein. Fibre-optic illumination of the sample

with LED light of 470 nm during the NMR experiments produces widespread changes in the backbone (NH) chemical shifts (comparison shown in Fig. 34).

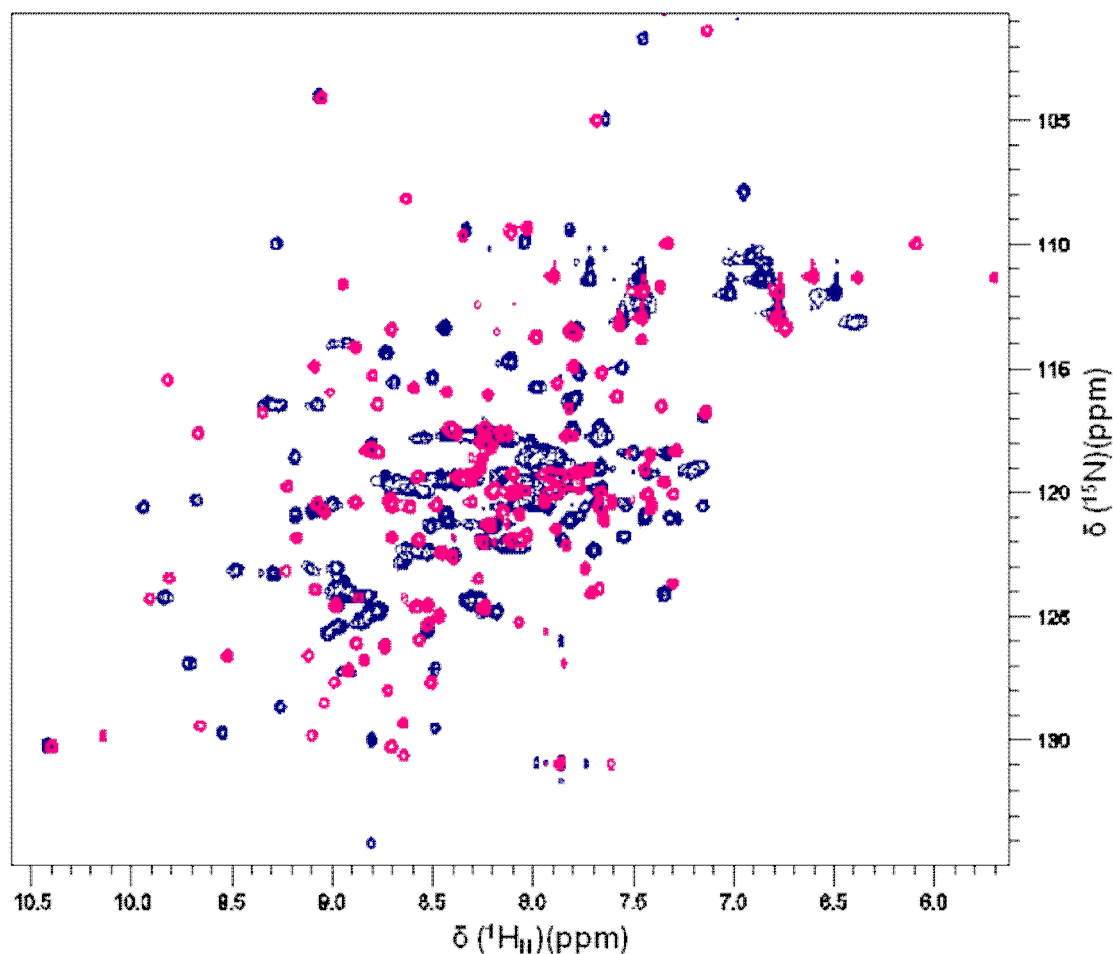


Figure 34: Overlay of ^1H - ^{15}N HSQC spectra of PpSB1-LOV recorded in the dark (blue) and after illumination with blue light (pink). All spectra were recorded at 40 °C at 900 MHz. Sample concentration was set to 0.5 mM using 10 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl and 1 mM DTT. Additionally, the sample contained 0.05 % (v/v) sodium azide as well as 10 % (v/v) D_2O .

Using a custom made fibre optics illumination device light-state NMR studies of PpSB1-LOV are possible. The comparison of ^1H - ^{15}N HSQC spectra of the protein in the dark- and light-state reveals widespread changes in the backbone amide chemical shifts indicative of a light-dependent structural rearrangement which has so far not been seen for PpSB1-LOV using other solution methods such as CD spectroscopy.

3.2.2.8 Sample integrity under NMR conditions

3.2.2.8.1 Photosensitivity and dark recovery under NMR conditions

As buffer conditions and the measuring temperature used in the here presented NMR experiments differ from the conditions under that the PpSB1-LOV protein has so far been studied, experiments were carried out to characterize the protein and ensure sample integrity under NMR measurement conditions. To study the LOV photocycle characteristics under NMR measurement conditions, UV/Vis spectra before and after illumination with blue light as well as dark recovery time traces were recorded (2.9.4.2) (Fig. 35). The UV/Vis spectra show typical changes associated with photocycling of LOV proteins (Fig. 35, panel A).

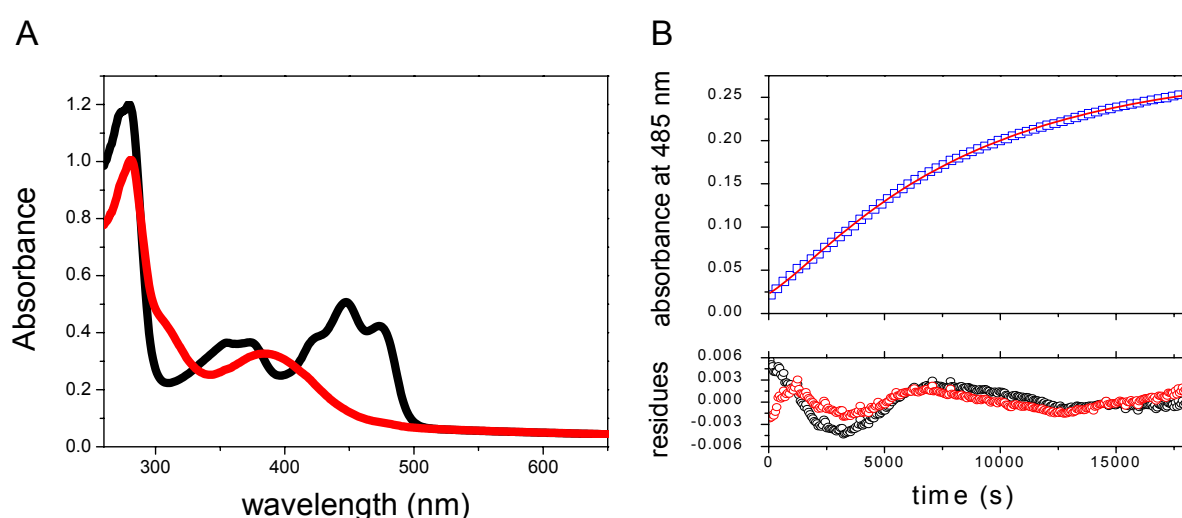


Figure 35: UV/Vis spectra (A) and dark recovery time trace recorded at 485 nm absorbance after blue-light illumination of PpSB1-LOV. (B) The lower panel displays the distribution of residuals for a single and double exponential fit of the experimental data, respectively. Residuals are better distributed for the double exponential fit. The average amplitude-weighted dark recovery time for PpSB1-LOV under the NMR measurement conditions was estimated to be $\tau_{\text{rec}} = 145$ min. The sample was diluted using the same buffer as for NMR experiments (10 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl, 1 mM DTT, 0.05% (v/v) sodium azide and 10% (v/v) D₂O). All measurements were carried out at 35 °C using a Beckmann DU650 UV/Vis spectrophotometer (2.9.4.2).

The recovery of the protein to the dark-state can be recorded by following the absorption at 485 nm after blue-light illumination (Fig. 35, panel B). Hereby, the kinetics of the dark recovery process correlate to the time the protein remains in its light-state (or signaling state) conformation even after light is switches off. This light-state lifetime is of primary importance for NMR studies conducted with illuminated LOV protein samples. LOV proteins that display very short light-state lifetimes show a rapid recovery of the protein to the dark-state when

light is switched off. Therefore, higher light intensities as well as continuous illumination is necessary to allow complete population of the light-state during NMR experiments.

After illumination a meta-stable light-state is formed which in the dark reverts with a time constant of $\tau_{\text{rec}} = 145$ minutes at 35 °C. In contrast at 20 °C in 10 mM sodium phosphate buffer pH 8.0, 10 mM NaCl, the protein displays a dark recovery time constant of about 2400 minutes. Thus in comparison to other LOV proteins that were studied by NMR such as the LOV2 domain of *Avena sativa* phototropin1 ($\tau_{\text{rec}} < 100$ s, at 20 °C, [76]), PpSB1-LOV displays a very stable and long-lived light-state. This allows the use of low light intensities for sample illumination and makes it easier to obtain well resolved light-state NMR spectra.

3.2.2.8.2 Determination of the native quaternary structure of PpSB1-LOV under NMR measurement conditions

Size-exclusion chromatography (SEC) was performed to elucidate the native oligomerization state of PpSB1-LOV under the same conditions as used in NMR experiments. Previous SEC studies have shown that in the dark PpSB1-LOV is predominately dimeric in solution [49]. However, those experiments were performed in a different buffer system at room-temperature. For comparison we here carried out SEC experiments using a buffer similar to the one used in NMR experiments (2.6.3.3). All experiments were performed employing a HPLC-SEC system (2.9.3.7) which allows heating of the column to the NMR measuring temperature of 40 °C. Buffers and samples were heated to 40 °C in a water bath. SEC runs were performed for the dark equilibrated PpSB1-LOV protein as well as for a sample illuminated with blue light. UV/Vis spectra of the respective HPLC-SEC elution peak were recorded to verify complete population of the PpSB1-LOV dark- or light-state, respectively (Fig. 36). The analysis revealed a predominately dimeric organization (about 90 %) in both the dark and light-state (Fig. 36). In both states about 10 % of an apparent monomeric species is detected. Moreover, UV/Vis spectra recorded for the two elution peaks (monomer/dimer) verify complete population of the respective state of the sample on the column.

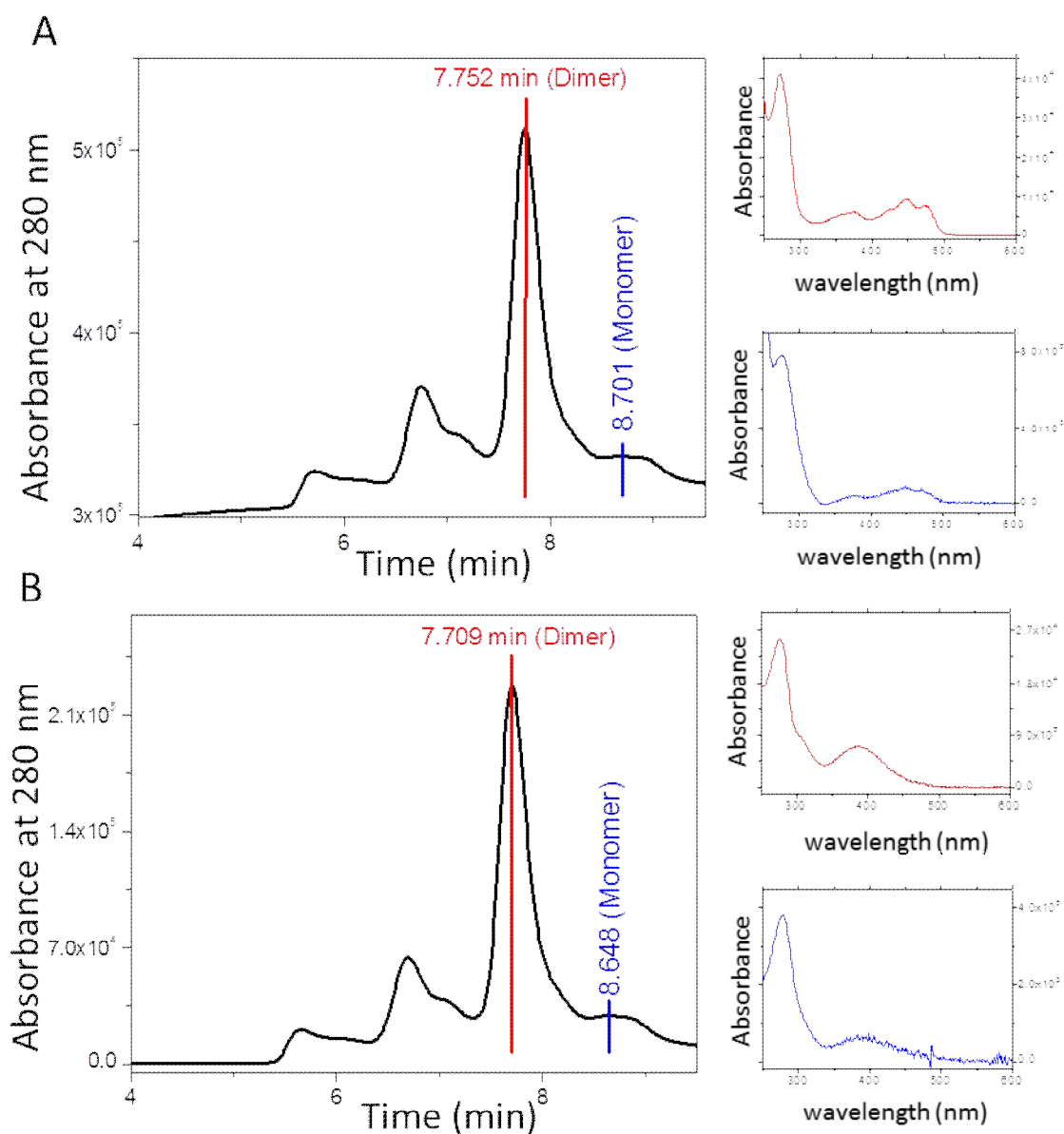


Figure 36: HPLC-SEC analysis of PpSB1-LOV in dark (A) and light-state (B). The SEC chromatogram is shown on the left side of the figure, while elution peak UV/Vis spectra are shown on the right side of both panels. The sample was run in 100 mM sodium phosphate buffer (pH 6.5) supplemented with 10 mM NaCl, 1 mM DTT. Buffer and column were heated to 40 °C.

Under the buffer and temperature conditions employed in the here presented NMR spectroscopic studies, PpSB1-LOV remains stable. UV/Vis spectroscopic studies reveal photocycling as well as a long-lived light state even at elevated temperatures. HPLC-SEC analysis hint at a predominately dimeric quaternary structure of the protein which appears to be independent of illumination.

3.2.3 Backbone resonance assignment of the PpSB1-LOV dark- and light-state

To facilitate assignment of all protein backbone resonances, a set of 2D and 3D NMR experiments were measured (see chapter 2.11.6.1). The assignment procedure used to assign dark- and light-state protein backbone resonances is outlined in the following.

Backbone resonances of PpSB1-LOV were assigned manually using the CcpNMR tool [106] employing sequential C_α , CO and NH chemical shift information derived from an array of triple resonance experiments (HNCA, HN(CO)CA, HNCO, HN(CA)CO) [96, 140]. All experiments contain ^{15}N - and ^1H - resonances, therefore allowing the use of this pair of spins as reference. The HNCA experiment correlates the ^1H and ^{15}N chemical shifts of residue (i) and $^{13}\text{C}_\alpha$ shifts of residue (i) and residue (i-1). Generally in HNCA, the intensity of the C_α (i) will be higher than the C_α (i-1). Thus the HN(CO)CA enables only an unambiguous assignment in case of accidental overlap of intra- and interresidue correlation. In order to establish the sequential assignment stretches of amino acid sequence were striped. The C_α chemical shift was compared with the shifts expected for particular amino acids. The expected chemical shifts for specific amino acids are shown in Fig. 37. Using this reference chemical shift list all amino acid types were assigned.

Representative strips of HNCA and HN(CO)CA spectra are shown in Fig. 38 which have been arranged in order of the sequence for clarity. The different spin systems are connected with each other by scalar coupling between heteronuclei in the protein backbone. Amino acids such as prolines are not observed due to lack of a backbone amide functional group.

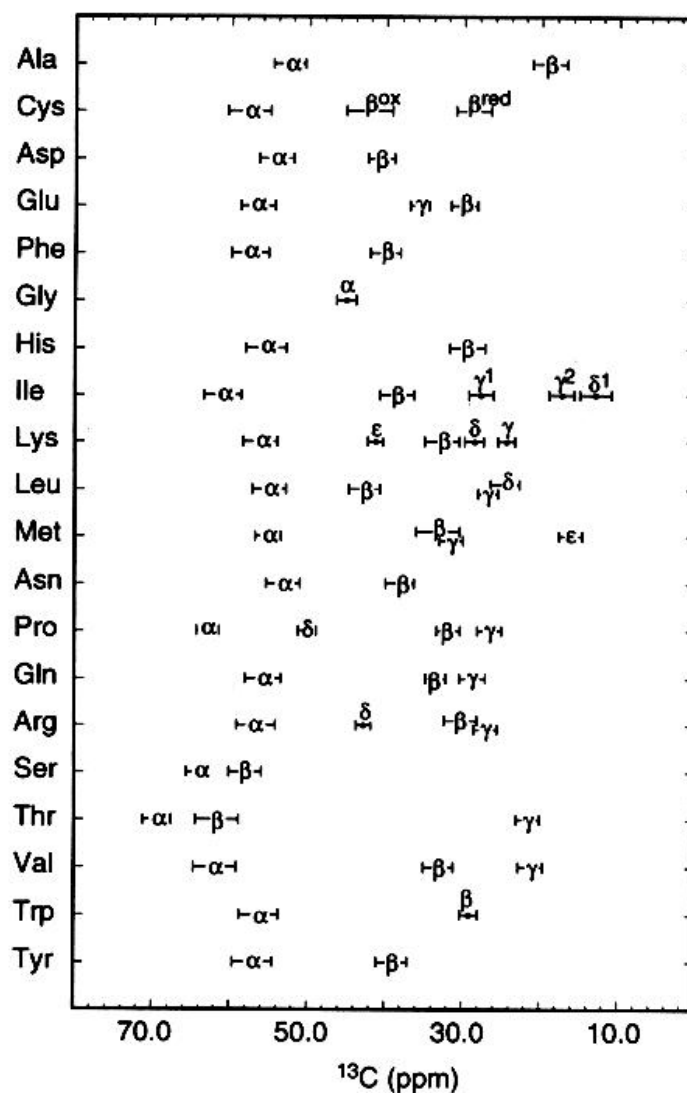


Figure 37: Average C_α and C_β chemical shifts for all 20 amino acids. Listed values were derived from a database of 13 proteins (Cavanagh *et al*, 1996).

The corresponding CO chemical shift values were extracted from HNCO and HN(CA)CO experiments respectively. The above strategy was used to assign all the amino acid backbone resonances of PpSB1-LOV in the dark state (Fig. 39) and the light state (Fig. 40).

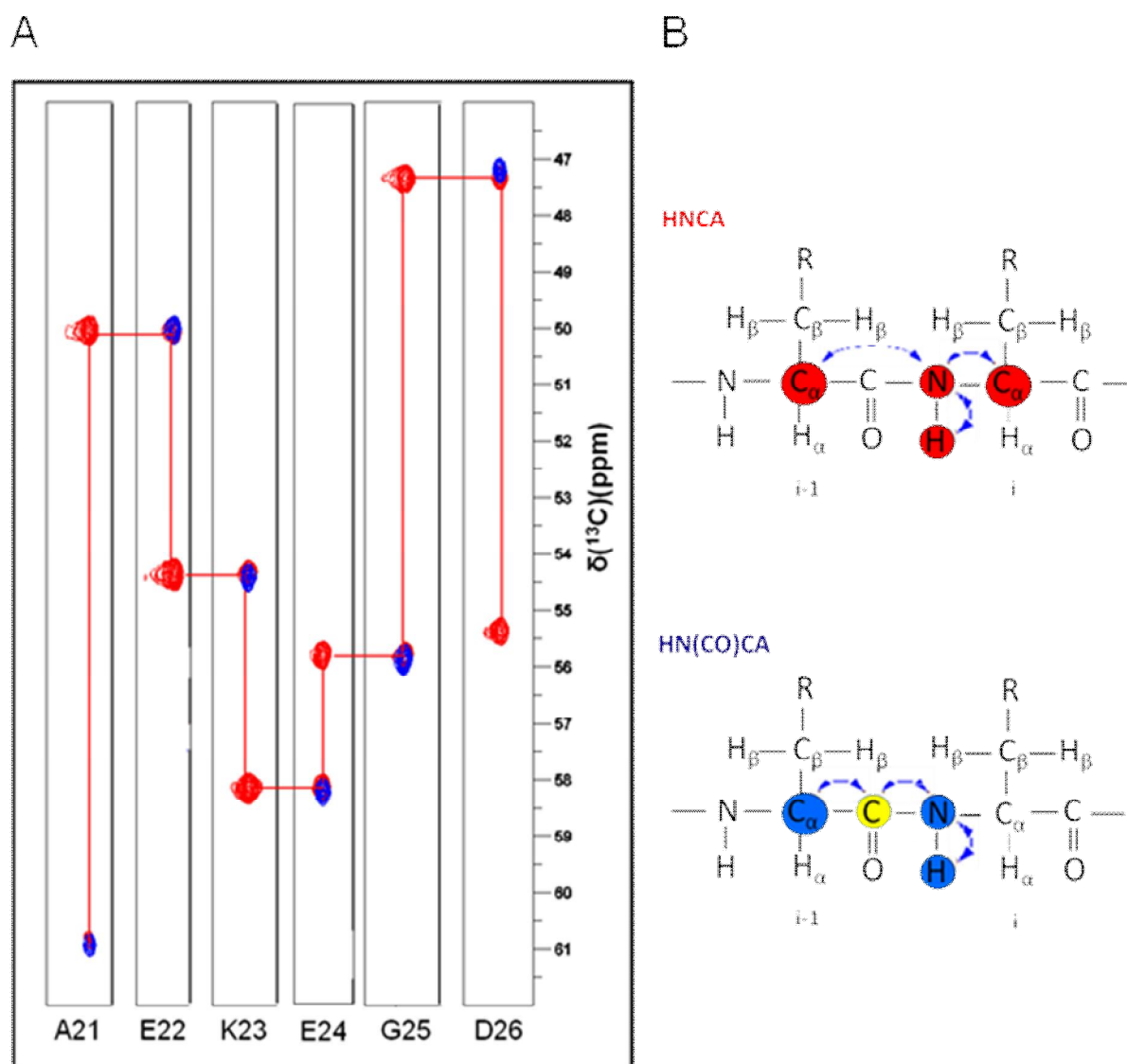


Figure 38: (A) Representative strips of HNCA and HN(CO)CA spectra, arranged in order of the sequence. The blue peaks represent the signal found in the HN(CO)CA spectra corresponding to the $i-1$ residue. The red peaks represent the signals found in HNCA spectra of residue i (higher intensity) and residue $i-1$ (lower intensity) (B) Schematic representation of the magnetization transfer utilized in HNCA and HN(CO)CA experiments. In HNCA experiments the magnetization is transferred from $^1\text{H}^{\text{N}}-^{15}\text{N}-\text{C}_\alpha(i)$ and $^1\text{H}^{\text{N}}-^{15}\text{N}-\text{C}_\alpha(i-1)$. In HN(CO)CA experiments magnetization is transferred from the $^1\text{H}^{\text{N}}-^{15}\text{N}-\text{C}_\alpha(i-1)$. Magnetization transfer is indicated by blue arrows.

3.2.3.1 Backbone assignment for the PpSB1-LOV dark-state

The backbone $^1\text{H}_\text{N}$, ^{15}N , ^{13}CO and $^{13}\text{C}_\alpha$ assignment for PpSB1-LOV in the dark-state has been completed to 99% of the visible backbone resonances (Fig. 39). No assignment was possible for four Pro residues (78P, 101P, 136P and 138P), seven residues within the N-terminal

region (residue 1M to residue 7L) and residue 64L because the corresponding resonances were not detectable in the respective NMR experiments.

Side-chain resonances of Asn and Gln (horizontal lines on the upper right corner) in ^1H - ^{15}N HSQC spectra have not been assigned for the dark state. For the light-state those resonances were assignable with the help of the assigned side-chain resonances from H(C)CH-TOCSY and (^1H - ^{13}C - ^1H)-HSQC-NOESY experiments. Due to time restrictions, those could not be assigned for the dark state.

3.2.3.2 Backbone and side-chain assignment for the light-state

The backbone ^1HN , ^{15}N , ^{13}CO and $^{13}\text{C}\alpha$ assignment of the PpSB1-LOV light-state has been completed to 99% of the visible backbone resonances (Fig. 40). No assignment was possible for four Pro residues (78P, 101P, 136P and 138P), four residues (1M, 2I, 3N and 4A) within the N-terminal region as well as for residue 64L which were not detectable in the respective NMR experiments.

For the light-state of PpSB1-LOV the side-chain proton assignment is 89% complete. Unassigned side-chain resonances mainly are such as $\text{H}\epsilon$ - $\text{H}\eta\text{b}^*$ of sixteen Arg residues, $\text{C}\epsilon/\text{H}\epsilon$ of Met residues, $\text{C}\delta$ in Gln, $\text{C}\gamma$ in Asn, $\text{C}\zeta$ in Arg, $\text{C}\delta$ in some Ile, $\text{C}\delta$ in Glu, $\text{C}\gamma$ in Asp and a number of signals due to Phe, Tyr and Trp ring signals.

Some additional, unassignable signals of weaker intensity are visible in light-state spectra, e.g. near to the 94W side-chain peak. Those signals however, were not visible in all experiments and could be due to either sample degradation (proteolysis) or might be related to loss or photodamage of the FMN chromophore during the experiment which in either case would result in a certain degree of structural heterogeneity of the sample.

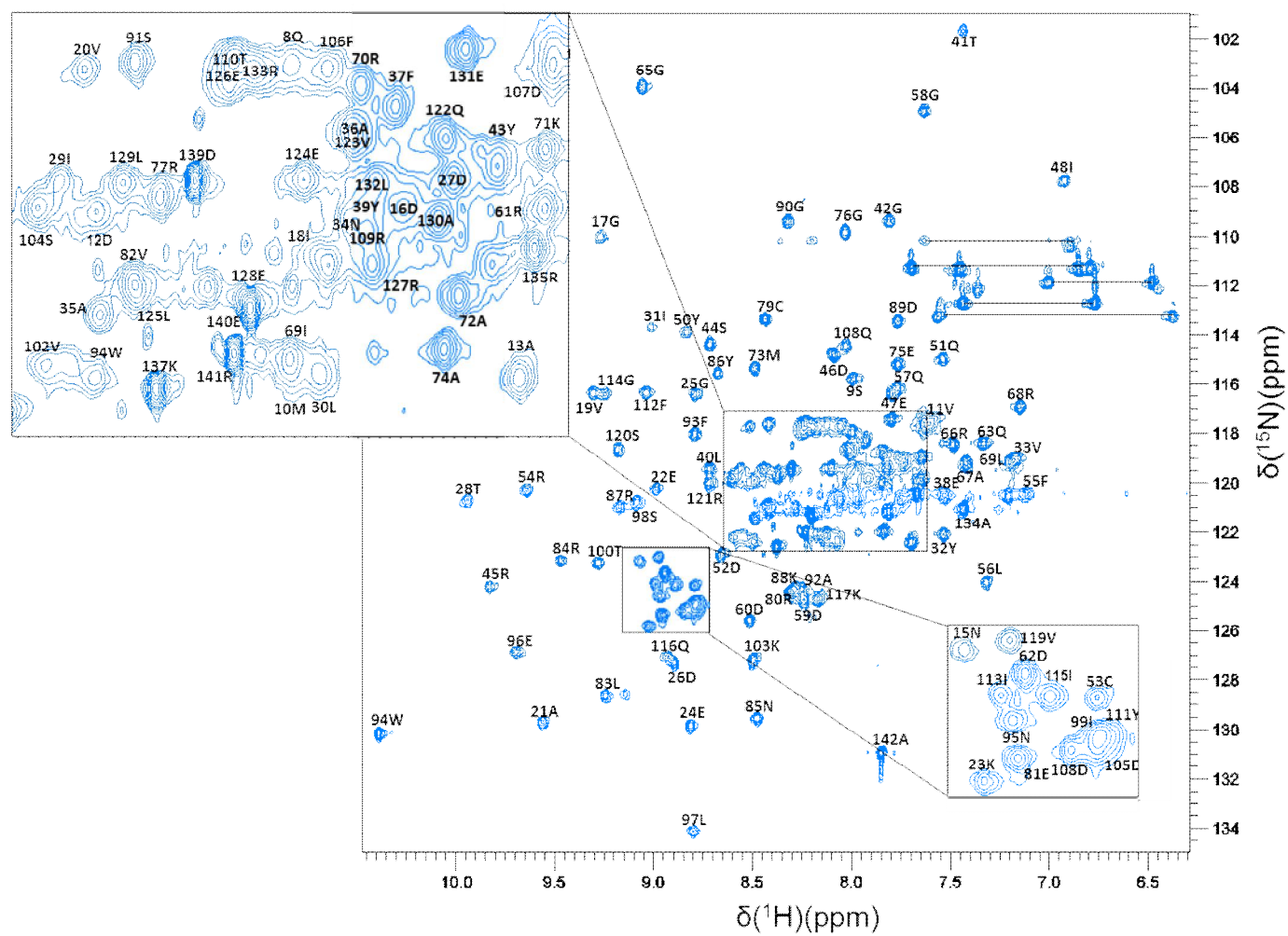


Figure 39: Assigned ^1H - ^{15}N HSQC spectrum of PpSB1-LOV in the dark-state. Backbone resonance residue assignments are indicated by *one-letter* amino acid code and the sequence number. Unassigned side-chain amide groups visible in ^1H - ^{15}N HSQC spectra are seen in upper right corner.

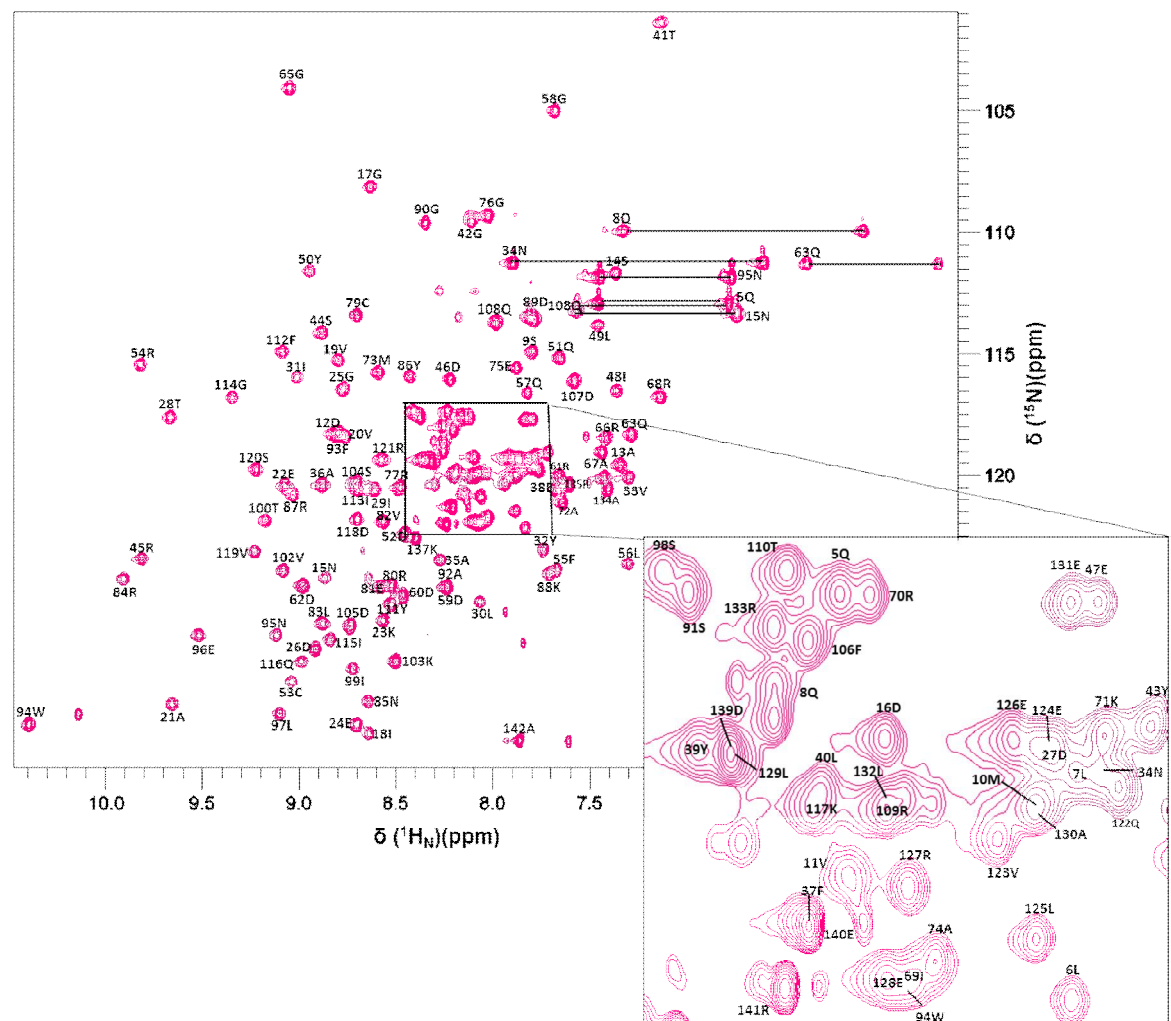


Figure 40: Assigned ^1H - ^{15}N -HSQC spectrum of PpSB1-LOV in the light-state. Backbone resonance assignments are indicated by *one-letter* aminoacid code and the sequence number. The side chain amide groups of asparagine and glutamine residues are connected by *horizontal lines*.

A series of double and triple resonance experiments allowed the near complete assignment of the protein backbone resonances of PpSB1-LOV in both the dark- and light-state. This assignment represents the first step necessary for atomic resolution structure calculations. Moreover, it enables a number of different NMR experiments which can reveal structural dynamics of the PpSB1-LOV protein in solution.

3.2.4 PpSB1-LOV secondary structure information obtained from NMR data

Based on amide backbone chemical shift values, protein residue-wise secondary structure can be predicted from NMR experiments using the TALOS+ software (Fig. 41). The prediction reveals a typical LOV core fold consisting of five β -strands and 4 α -helices in the topological order A β , B β , C α , D α , E α , F α , G β , H β and I β . Outside the conserved LOV core, a short N-terminal cap containing the helix A' α (highlighted in light green) and a longer C-terminal J α -helix extension (highlighted in dark blue) can be identified. For comparison, the corresponding secondary structure elements obtained from the pdb coordinates of AsLOV2 (pdb-entry: 2V1A) and YtvA-LOV (pdb-entry: 2PR5) are shown in Fig. 41.

The most pronounced difference in secondary structure between the PpSB1-LOV dark- and light state is apparent for the region of the B β -strand. While for the light-state an extended β -strand conformation is proposed by TALOS+, dark-state data imply loss of secondary structure for the respective region.

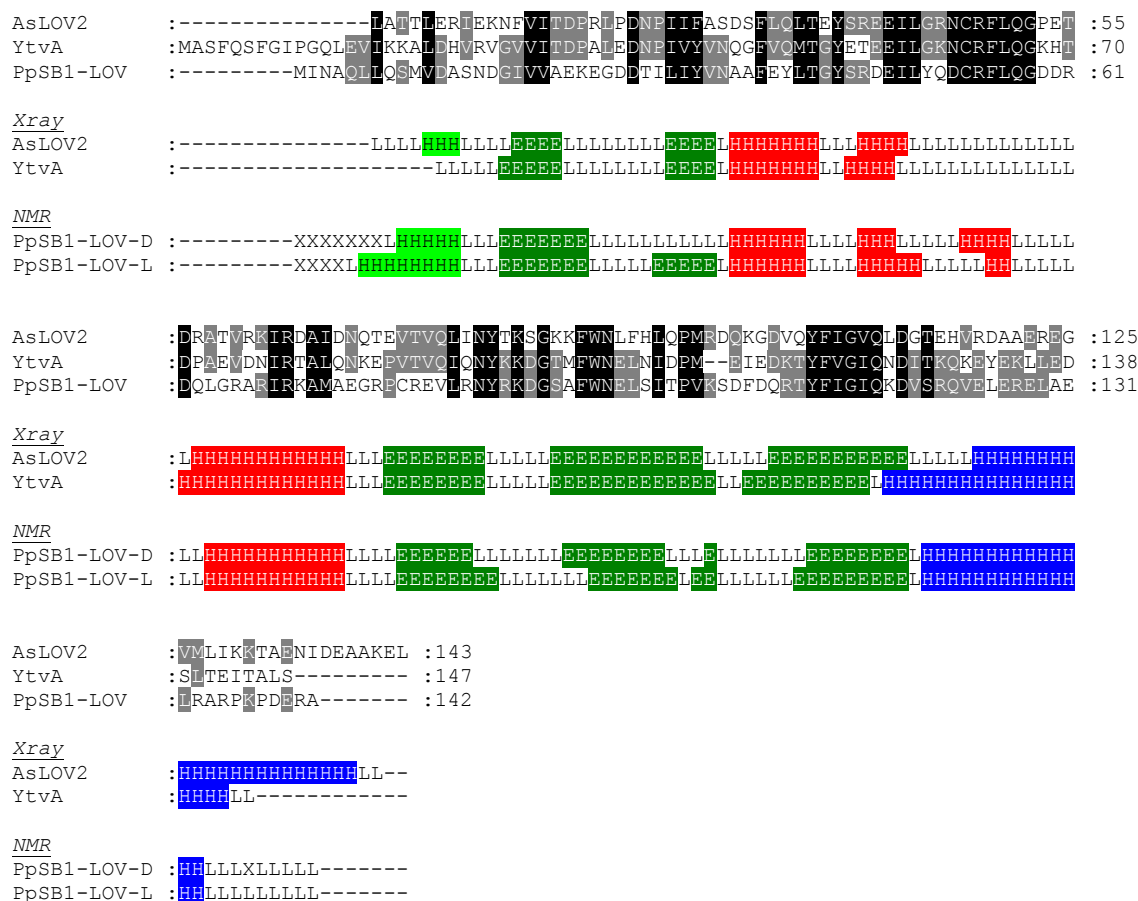


Figure 41: Multiple sequence alignment illustrating the NMR (TALOS+) derived secondary structure content of PpSB1-LOV in the dark- (PpSB1-LOV-D) and light-state (PpSB1-LOV-L) For comparison the same information is shown for oat phototropin1 LOV2 domain (AsLOV2) and the *B. subtilis* YtvA-LOV domain (YtvA). For the latter two proteins secondary structure information was obtained from the respective pdb coordinates (2V1A and 2PR5, respectively).

NMR-based secondary structure predictions suggest a canonical LOV/PAS fold for PpSB1-LOV and reveal the presence of auxiliary α -helical elements N- and C-terminally of the conserved LOV core. Although care has to be taken, when interpreting TALOS+ secondary structure predictions, the analysis nevertheless hints at local changes in secondary structure associated with illumination of PpSB1-LOV.

3.2.5 PpSB1-LOV structural dynamics studied by NMR

Based on the backbone resonance assignment for the dark- and light state of PpSB1-LOV additional NMR experiments are possible that can reveal the dynamic behavior of the studied protein. Such experiments include T1 and T2 relaxation studies as well as hydrogen-deuterium (H/D) exchange measurements (see 1.3.5.3 in the introduction for details). The results of those studies will be presented in the following chapters.

3.2.5.1 Dynamics of PpSB1-LOV in the dark-state

^1H - ^{15}N NOE, ^{15}N longitudinal (T1) and transverse (T2) relaxation experiments were measured at 40 °C using a ^{15}N -labelled PpSB1-LOV sample (see 1.3.5.5 and 2.11.1 for details). All experiments were conducted for dark-equilibrated and light-state PpSB1-LOV samples. Fig. 42 (panel A) depicts the quotient of longitudinal (R1) and transverse (R2) ^{15}N -relaxation rates (R1/R2) for the dark-state of PpSB1-LOV. The corresponding ^1H - ^{15}N hetNOEs are shown in Fig. 42 (panel B). The R1/R2 and hetNOE values are evenly distributed throughout the whole structured part of the protein (from A' α to J α) indicative of a well folded rather rigid protein domain. At the immediate C-terminal end (after J α) R1/R2 values are increased and the corresponding hetNOEs drop to negative values which is indicative of high structural flexibility.

The respective values for the light-state are shown in Fig. 43. As for the dark-state R1/R2 values are evenly distributed throughout most of the PpSB1-LOV protein. A slight increase in R1/R2 and concomitant decrease in hetNOE values compared to the dark-state is, at the first glance, only apparent for the H β -I β loop. At the immediate C-terminal end R1/R2 values are increased and hetNOE values drop below zero. Hence the C-terminal end appears to be equally flexible in both dark- and light-state.

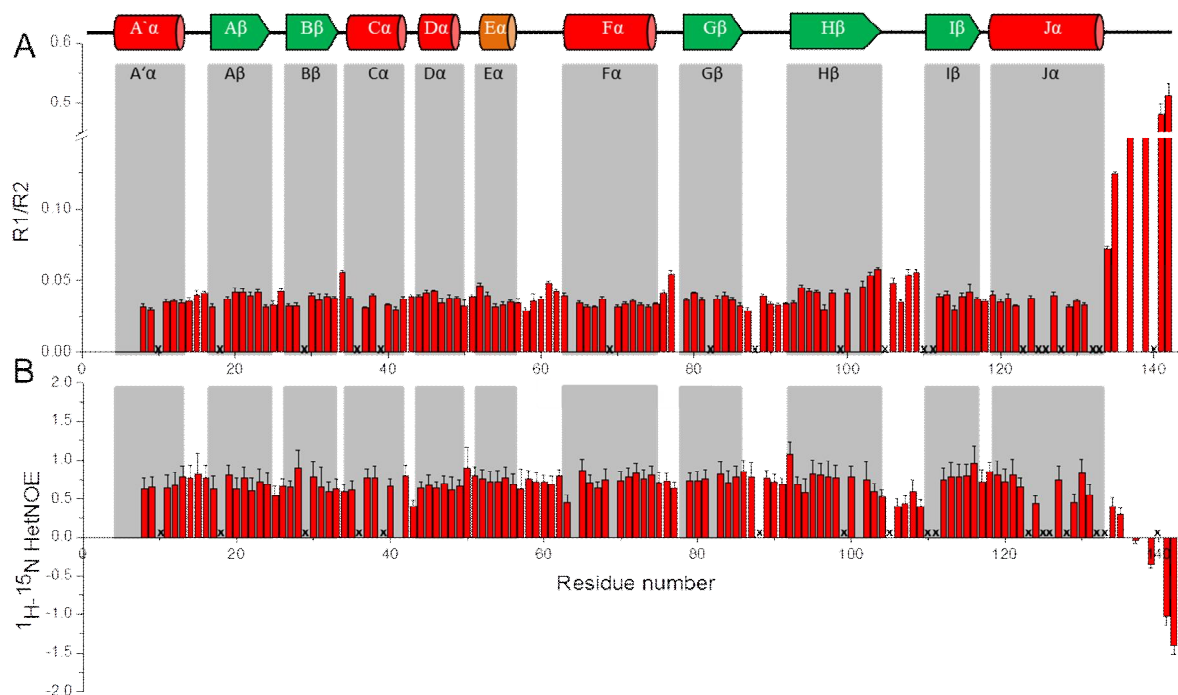


Figure 42: Relaxation data for PpSB1-LOV in the dark state. Quotient of longitudinal (R1) and transverse (R2) ^{15}N relaxation rates (R1/R2) (**A**) and ^1H - ^{15}N hetNOEs (**B**) are plotted against residue number. Missing residues are shown with open bars. Residues showing signal overlap which made proper peak integration impossible are marked by a cross. Secondary structure elements characteristic for LOV domains are shown above the plot.

Residues that were exchange-broadened beyond detection in all light-state T1, T2 and hetNOEs measurements are Cys53 and Gly17 (marked in Fig. 43). Unfortunately, line-broadening effects and also spectral overlap prevented analysis of the motions in more detail.

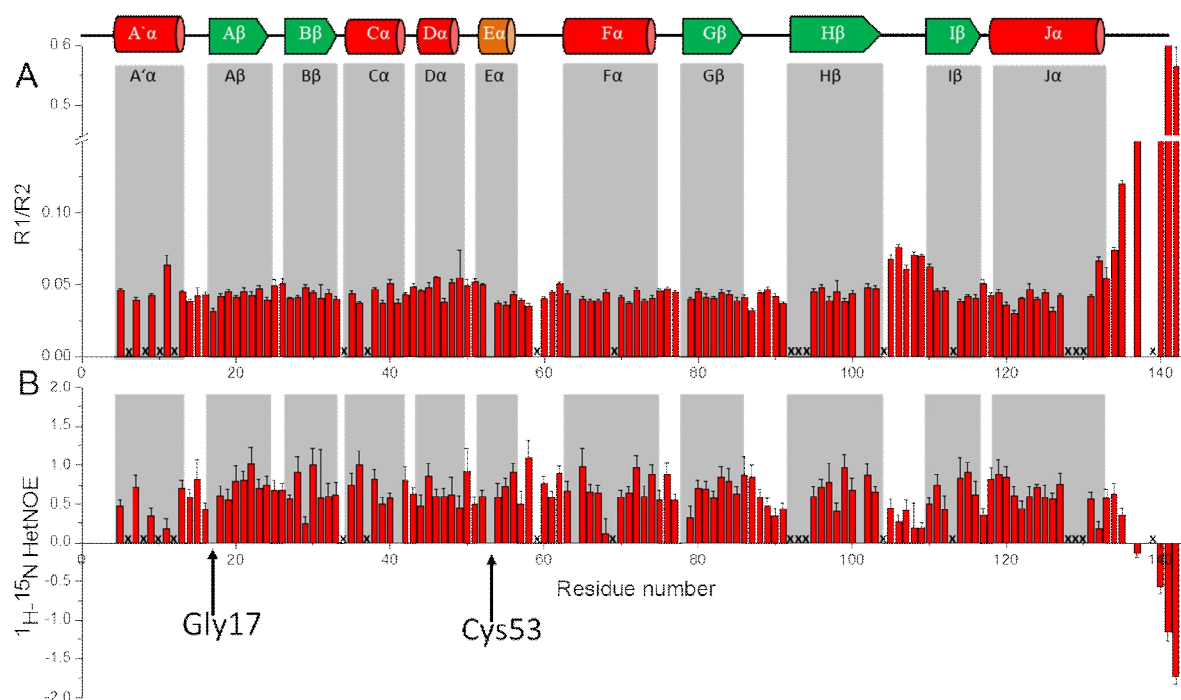


Figure 43: Relaxation data for PpSB1-LOV in the light state. Quotient of longitudinal (R1) and transverse (R2) ^{15}N relaxation rates (R1/R2) (A) and ^1H - ^{15}N hetNOEs (B) are plotted against residue number. Missing residues are shown with open bars. Residues with signal overlap for which no proper peak integration was possible are marked by a cross. Secondary structure elements characteristic for LOV domains are shown above the plot. The photoactive cystein residue (C53) of PpSB1-LOV which forms a covalent bond with the C4a atom of the flavin isoalloxazine ring is highlighted.

While longitudinal (R1) and transverse (R2) relaxation rates are very similar in the dark- and light-state, disappearance of the backbone resonance peak for the photoactive C53 residue is only apparent in light-state experiments. This might be due to covalent bond formation between the thiol group of Cys53 and the C4a atom of the flavin isoalloxazine ring in the light state of the protein, providing further evidence for complete population of the PpSB1-LOV light-state in the NMR tube.

3.2.5.2 Dark and light state solvent accessibility studied by proton-deuterium (H/D) exchange experiment

H/D exchange experiments can provide a quantitative measure for the solvent accessibility of a given amino acid in the protein (see introduction 1.3.5.3 and materials and methods 2.11.9).

Hereby, the H to D exchange rate is slower when the amide proton is buried in the core of the protein or when it is involved in an H-bond. Consequently the rate is faster when the amide proton is exposed to solvent. The H/D exchange of amide protons in protein backbones has become an important way to address the dynamics of a protein in solution [156]. To address the solvent accessibility of the PpSB1-LOV in the dark- and light-state, an H/D exchange experiment was conducted using ^{15}N -labelled PpSB1-LOV protein (see Materials and Methods 2.11.9 for details).

The cross peaks of some residues directly disappeared in the first HSQC frame after D_2O was added, which indicated that these residues are in regions of high flexibility or solvent accessibility. Data quality did not allow quantitative decay rate analysis therefore H/D exchange rates were sorted in distinct groups. Residues for which the H/D exchange reaction was almost completed after 300 s (first HSQC spectrum after addition of D_2O) are shown in gray. Those represent highly flexible or solvent exposed residues. For the remaining residues exchange rates could be estimated from the series of HSQC spectra. Because of data quality exchange rates were sorted into distinct groups: I) residues with an exchange rate below $1.0\text{E}-2$, II) residues with an exchange rate below $2.5\text{E}-3$ and III) residues with an exchange rate below $3.4\text{E}-4$ and IV) residues with an exchange rate below $3.0\text{E}-5$.

Residues showing a rapid H/D exchange rate in the dark-state include the first N-terminal 10 residues (including parts of $A'\alpha$), most loop regions, the helices $D\alpha$, $E\alpha$ and $F\alpha$ as well as the complete C-terminal $J\alpha$ -helix as well as all ensuing C-terminal residues. Differences in the respective light-state H/D exchange experiments are most prominent for the helices $D\alpha$ and $E\alpha$, which exchange rapidly in the dark (no detectable signal in first HSQC spectrum) but display markedly slower exchange rates in the light-state (Fig. 44, panel B). In both states, pronounced exchange protected residues are mostly located within β -strands of the central LOV β -scaffold.

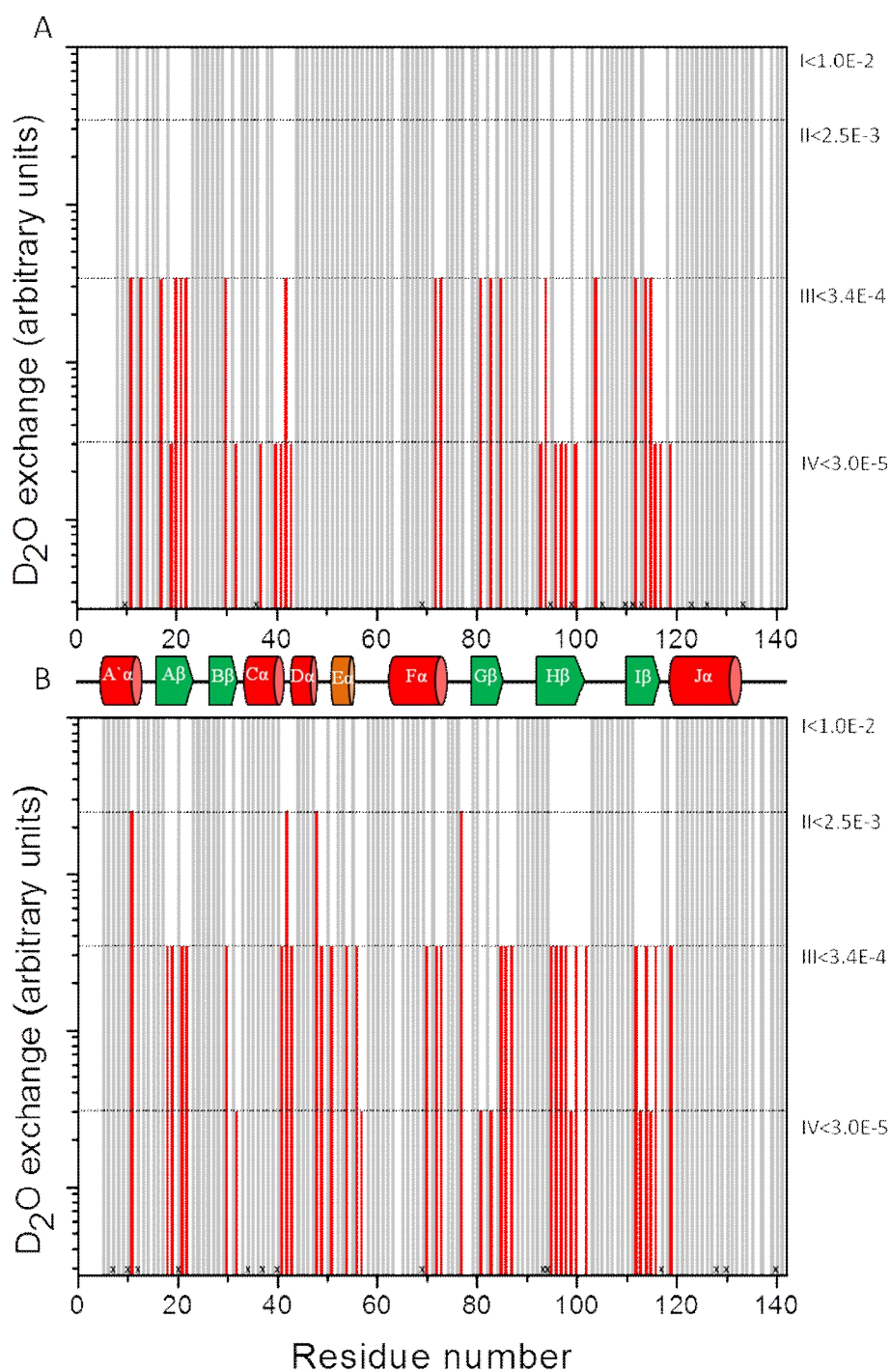


Figure 44: H/D exchange in the dark- (**A**) and light-state (**B**) of PpSB1-LOV. Residues for which the H/D exchange reaction was almost completed after 300 s are shown in gray. Open bars represent residues that were not assigned in the HSQC spectrum (see before). Residues where spectral overlap did not allow proper peak integration are marked by a cross. For the remaining residues exchange rates could be estimated from the series of HSQC spectra. Because of data quality exchange rates were sorted into distinct groups: I) residues with an exchange rate below $1.0E-2$, II) residues with an exchange rate below $2.5E-3$ and III) residues with an exchange rate below $3.4E-4$ and IV) residues with an exchange rate below $3.0E-5$.

H/D exchange experiments provide information related to protected (buried) regions of PpSB1-LOV. In the dark-state, mostly residues of the central β -scaffold appear protected against deuterium exchange. Upon illumination of the protein additionally residues on $D\alpha$ and $E\alpha$ become more protected. H/D exchange data for the C-terminal $J\alpha$ -helix suggest that the respective region is solvent exposed in both the dark- and light state. Likewise, the N-terminal capping helix $A'\alpha$ appears equally solvent exposed in both states.

Chapter 4: Discussion

4.1. The PpSB1-LOV light-state X-ray structure probed in solution

4.1.1 Structural features of the recently solved PpSB1-LOV X-ray structure

As outlined in the introduction (1.2.3.1) the two homologous short LOV proteins PpSB1-LOV and PpSB2-LOV from *P. putida* KT2440 lack a fused effector domain, but instead contain short N- and C-terminal extensions outside the conserved LOV core (see introduction 1.2.3.1 and chapter 3.1 of the results section). Recent studies on other LOV photosensory systems highlighted the importance of those latter auxiliary structural elements for the intramolecular signal-relay process [72, 73, 76].

At the time this study was initiated, no X-ray structure of either PpSB1-LOV or PpSB2-LOV was available. Hence, i.e. the structure and orientation of the N- and C-terminal extensions as well as atomic resolution details about the LOV core domain structure were unknown. This changed with the recently presented light-state structure of PpSB1-LOV which was obtained by X-ray crystallography [157].

For PpSB1-LOV well diffracting crystals could only be obtained for the protein crystallized under constant low-level white-light illumination [157]. Under the same crystallization conditions, no crystals formed in the dark. Moreover, single crystal micro-spectrometry data suggested that the protein is trapped in the light-state within the crystal. While flavin-C4a *sp*³ hybridization as well as electron density in the region of the covalent FMN-Cys53 bond was visible in the σ -A weighted electron density ($2mFo - DFc$) map, the observed FMN-C4a \cdots Cys-S γ distance was with 2.35 Å found to be too long for a covalent bond but too short for a van der Waals contact. Bond disruption most probably occurred due to radiation damage imposed by the X-ray beam during the diffraction measurement. Hence, the structure was interpreted as light-state showing some degree of radiation damage and thus essentially represents a “mixed state with respect to photo-activation” [157].

While this study provided a first glimpse into atomistic details of the FMN-protein interaction as well as into the structural nature of the PpSB1-LOV protein in the light-state, it yet failed

to provide any information about the mechanism of signal-relay and photo-activation. This is due to the fact that no dark-state structure of PpSB1-LOV is available and hence any information regarding the light-induced conformational changes of the protein remains inaccessible. These observations stress the importance of further structural studies on the PpSB1-LOV photoreceptor protein by NMR spectroscopy. Fig. 45 depicts the light-state structure of PpSB1-LOV. In the crystal, two PpSB1-LOV molecules are found in the asymmetric unit suggesting a dimeric organization of the protein. Interestingly the N-terminal cap ($A'\alpha$) and the C-terminal $J\alpha$ -helix protrude from the LOV core and support dimerization of the protein (Fig. 45).

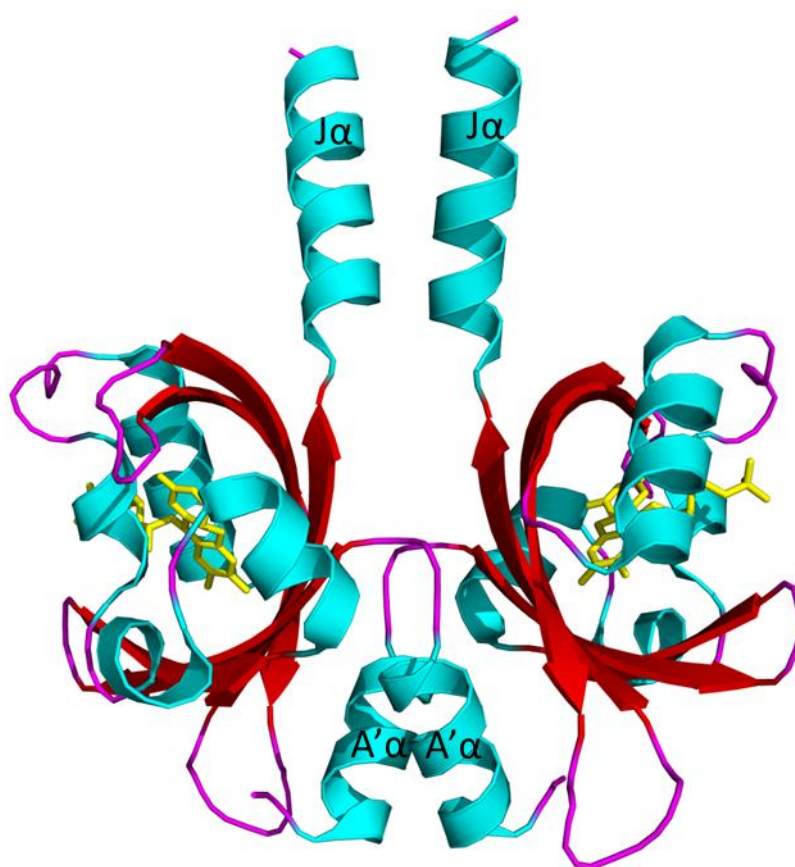


Figure 45: Light-state X-ray structure of PpSB1-LOV [157] (pdb entry: 3SW1). Apart from the structurally highly conserved LOV core the protein comprises protruding N-terminal ($A'\alpha$) and C-terminal ($J\alpha$) α -helical extensions (labeled in black). α -helices are shown in cyan, β -strands in red and random coiled secondary structures in pink. Two PpSB1-LOV molecules are found in the asymmetric unit. In each subunit of the dimeric protein one FMN molecule is bound (yellow sticks).

4.1.2 Probing the role of the N- and C-terminal extensions for the structural integrity of the *P. putida* LOV proteins

From X-ray crystallographic studies it is many times not clear whether the in the crystal observed quaternary structure is a result of crystal packing or represents the biologically relevant native oligomerization state of the protein in solution within the cell. Therefore, in the first part of this thesis we studied, the quaternary structure of PpSB1-LOV in solution (3.2), generated various truncated protein variants (3.1) and present CD and NMR spectroscopic data for the C-terminal part of the two short *P. putida* LOV proteins (3.1.3).

4.1.2.1 PpSB1-LOV appears to be dimeric in both the dark- and light-state

In order to verify dimerization of PpSB1-LOV in both dark- and light state we used a HPLC-based SEC system which facilitates the recording of elution peak UV/Vis spectra and thus enables unequivocal assignment of the “structural” state of the protein in solution on the SEC-column (see chapter 3.2.2.8.2). Those studies suggest that PpSB1-LOV is a dimer in solution in both the dark- and light-state. Moreover, dimerization is retained at higher temperature (40 °C) and lower pH values (pH 6.5) i.e. under the conditions similar to the ones used in NMR spectroscopic studies (chapter 3.2.2.8.2).

4.1.2.2 The helical N- and C-terminal extensions of PpSB1-LOV and PpSB2-LOV are essential for folding and structural integrity of both proteins

In order to study the role, structure and function of the N- and C-terminal extensions in solution several truncated PpSB1-LOV, and for comparison PpSB2-LOV, constructs were generated (chapter 3.1). Those studies revealed that the two structural elements are essential for the production of soluble protein in *E. coli*. The truncation of both the N- and C-terminal extensions resulted in a complete loss of protein solubility when expressed under similar conditions as the respective wild-type proteins (Fig. 11, chapter 3.1.1). This indicates that both regions contribute to the proper folding and structural integrity of both proteins *in vivo*. This is in contrast to multi-domain LOV photoreceptors, where both structural elements can be readily removed without loss of protein solubility [73, 101, 158].

In light of those findings, it seems likely that the observed dimer packing found in the PpSB1-LOV crystal is also observed in solution. Further evidence regarding the mode of subunit association will be discussed in chapter 4.2.2.

However, as a certain fraction of an apparent monomeric species (~10 %) is also observed in SEC analyses (3.2.2.8.2 or Fig. 36), we cannot rule out completely that this equilibrium might be shifted under the right experimental conditions.

Unfortunately, no soluble protein of the truncated PpSB1-LOV and PpSB2-LOV variants could be obtained. Therefore, no structural studies on those variants i.e. by CD- and NMR-spectroscopy were possible. Hence we had to resort to other means to analyze the folding of those structural elements in solution.

4.1.2.3 The isolated C-terminal extension of PpSB1-LOV and PpSB2-LOV adopt their native secondary structure independent from the rest of the protein

Another interesting feature of the PpSB1-LOV structure is the mode of $J\alpha$ -helix interaction in the dimer. The X-ray structure (Fig. 45) as well as here presented bioinformatic analyses (chapter 3.1.2) suggest that the protruding $J\alpha$ -extensions form a coiled-coil quaternary structural element much like it was suggested for *B. subtilis* YtvA [73] and several PAS-domain based chemosensors, whose sensor domains show structural similarity to LOV domains [159-164]. Coiled coils are widespread, independently folding interaction motifs present in many structurally different proteins [165]. Therefore, we here asked the question whether the $J\alpha$ -helix extension of PpSB1-LOV and PpSB2-LOV can fold independently from the rest of the protein. Given the short length of the $J\alpha$ -helices, we obtained both protein segments (PpSB1- $J\alpha$ and PpSB2- $J\alpha$) as custom synthesized N-terminally acetylated peptides.

CD spectra (Fig. 15, chapter 3.1.3.2) as well as the here presented solution NMR data (Fig. 19, chapter 3.1.3.5) indicate formation of α -helical secondary structure for the isolated PpSB1- $J\alpha$ and PpSB2- $J\alpha$ peptides at low temperature (CD spectroscopy, chapter 3.1.3.2) or in the presence of minimal amounts of the helix-stabilizing agent TFE (NMR spectroscopy, chapter 3.1.3.3). While the two isolated $J\alpha$ -elements seem to fold readily into a partially α -helical conformation in solution, direct spectral evidence for the formation of dimeric parallel coiled coils is missing. Both peptide CD spectra recorded at 1 °C show minima at

222 nm and 208 nm as well as a maximum at around 190 nm. These are signature features of helical secondary structure representing the $n \rightarrow \pi^*$ transition (~ 222 nm) and $\pi \rightarrow \pi^*$ exciton split transition (~ 190 nm and ~ 208 nm) of peptide bonds [166]. The ellipticity ratio $[\Theta_{222}]/[\Theta_{208}]$ has previously been used as an indicator for coiled-coil formation of peptides [167, 168]. Hereby, non-coiled helices show a ratio of about 0.8, whereas coiled-coil structures yield values close to 1 [167]. Mechanistically this is due to the fact that in coiled-coils the parallel polarized amide $\pi \rightarrow \pi^*$ transition becomes less dichroic, which effectively reduces the negativity of $[\Theta_{208}]$. Because $[\Theta_{222}]$ is not affected, the $[\Theta_{222}]/[\Theta_{208}]$ ratio increases to about 1. For PpSB1-J α and PpSB2-J α $[\Theta_{222}]/[\Theta_{208}]$ ratios are ~ 0.67 and ~ 0.89 . Thus it appears that anchoring of the J α -helix to the dimeric LOV-core domain is necessary to promote coiled-coil formation. The superimposition of the NMR derived peptide structure of PpSB1-J α (Fig. 46) to the respective structural element in the full-length crystal structure (pdb entry: 3SW1, [157]) yields backbone and all-atom RMSD values of 0.362 Å and 0.919 Å, respectively. To evaluate PpSB1-J α dimer formation, the energy-minimized NMR structure was aligned separately to both chains of the PpSB1-LOV X-ray structure. The segment corresponding to the respective PpSB1-J α dimer was retrieved. The superposition of the corresponding dimer model (dark grey) and the PpSB1-LOV X-ray structure (transparent, white) is shown in Fig. 46 (panel A). Moreover, a similar orientation of *a-d* heptad hydrophobic residues is observed in the NMR structure and in the crystal.

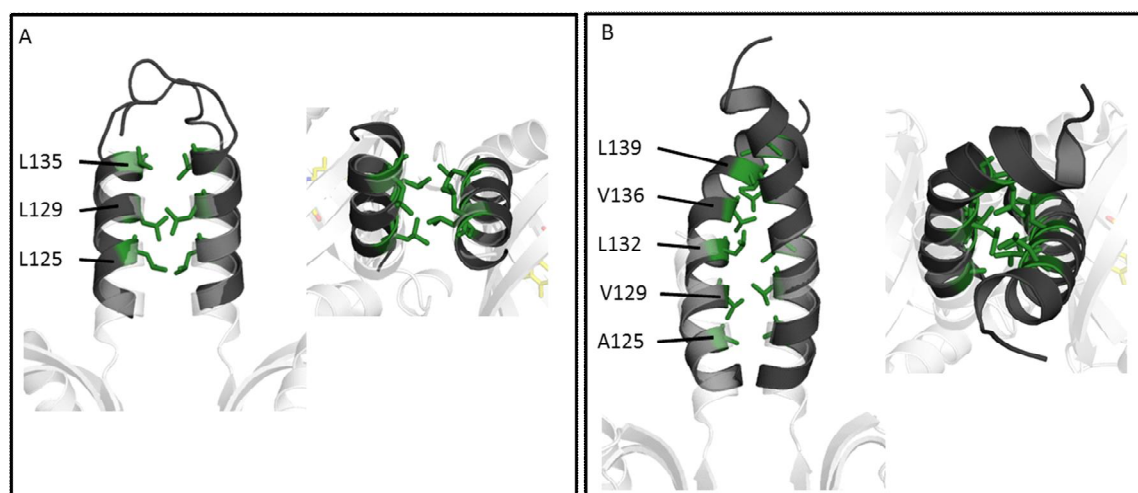


Figure 46: Superposition of the PpSB1-LOV X-ray structure (pdb entry: 3SW1) and the NMR derived PpSB1- $J\alpha$ (A) and PpSB2- $J\alpha$ NMR structure (B). For both NMR derived structures the hypothetical dimer obtained from docking studies is depicted (dark grey). The PpSB1-LOV X-ray structure is shown as transparent white cartoon drawing. The PpSB1- $J\alpha$ dimer model was obtained by superimposing the monomeric NMR derived PpSB1- $J\alpha$ structure on the individual chains of the PpSB1-LOV X-ray structure. The PpSB2- $J\alpha$ model was obtained using the ClusPro 2.0 webserver (<http://cluspro.bu.edu>) [169]. The “best” ClusPro model was selected based on favorable superposition onto the PpSB1-LOV X-ray structure. Buried hydrophobic residues comprising the residues *a-d* of a coiled-coil heptad repeat are highlighted in green (stick representation). For better visibility, the structure was rotated and the respective coiled-coil segment enlarged (right side of each panel).

In order to evaluate the possibility of dimer formation for the PpSB2- $J\alpha$ peptide, the energy minimized PpSB2- $J\alpha$ monomer NMR structure was subjected to docking studies. For homomultimer docking the ClusPro 2.0 webserver (<http://cluspro.bu.edu>) was used [169]. The “best” ClusPro model was selected considering a favorable superposition (lowest backbone RMSD) onto the $J\alpha$ -helix segment of the PpSB1-LOV dimer structure. Fig. 46 (panel B) depicts the superposition of the “best” PpSB2- $J\alpha$ dimer model (dark grey) and the PpSB1-LOV X-ray structure (transparent, white). Hydrophobic residues that are buried by coiled-coil formation, representing the *a-d* positions of a canonical *abcdefg* heptad-repeat (see also Fig. 13), are highlighted in green and the respective side-chains are shown in stick representation. Thus, while the general folding is conserved in both crystal and solution, the $J\alpha$ structural elements possess apparently a certain degree of quaternary structural conformational flexibility. They either exist as isolated helix in solution or packed as a tight coiled coil in the context of the full-length protein in the crystal. Of course, one possibility is that coiled-coil formation observed in the crystal could be an artifact related to crystal packing. While the here presented data on truncated PpSB1-LOV/PpSB2-LOV protein constructs (chapter 3.1.1) as well as size-exclusion chromatographic data (chapter 3.2.2.8.2) argues against this

scenario, the issue could only be resolved unequivocally by obtaining high-resolution NMR structures of the PpSB1-LOV protein. More likely, the conformational flexibility observed for the $J\alpha$ -helix element in solution and the crystal structure could be related to coiled-coil instability resulting from the rather short sequence length of the PpSB1-LOV/PpSB2-LOV $J\alpha$ -helix. Thus, in conclusion it appears that the $J\alpha$ -helix coiled-coil observed in the crystal is not sufficiently stable to promote coiled-coil formation in solution in the absence of the LOV-core domain. Anchoring of the $J\alpha$ -helix at the C-terminus of the LOV core in the full-length protein might impose dynamic- and distance restraints on the orientation of the $J\alpha$ -helix and thus promote coiled-coil formation as observed in the crystal structure. Given, those structural features, the C-terminal $J\alpha$ -helix could well represent the “missing” effector domain of the *P. putida* short LOV protein family whose interaction with downstream signaling partners could be modulated by blue light (further discussed in chapter 4.6).

4.2 Light-state NMR studies hint at a grossly similar structure of PpSB1-LOV in solution and the crystal

Using the near complete backbone resonance assignment (3.2.3) and the partial side-chain assignment information (3.2.3.3) obtained for the light-state of PpSB1-LOV, several analyses can be performed to study the overall secondary structure of the protein in solution (see chapter 3.2.4).

4.2.1 Comparison of NMR-based secondary structure predictions for PpSB1-LOV in the light-state and the crystal structure

Based on backbone chemical shift values, protein residue-wise secondary structure can be predicted from NMR experiments using the TALOS+ software [114]. For the light-state of PpSB1-LOV the prediction reveals a typical LOV core fold consisting of five β -strands and 4 α -helices in the topological order $A\beta$, $B\beta$, $C\alpha$, $D\alpha$, $E\alpha$, $F\alpha$, $G\beta$, $H\beta$ and $I\beta$. Outside the conserved LOV core, a short N-terminal cap containing the helix $A'\alpha$ and a longer C-terminal $J\alpha$ -helix extension can be identified. For the sequence region between Tyr50 and Gln63 TALOS+ predicts only two residues in helical conformation. This is in line with the recently solved light-state X-ray structure where this region comprises a short 3_{10} -helix ($E\alpha$). Globally, the NMR derived secondary structure for the light state of PpSB1-LOV in solution is in good agreement with the secondary structure derived for the PpSB1-LOV light-state X-

ray structure [157] (Table 13). The deviation between the NMR (TALOS+, [114]) derived secondary structure content and the secondary structure content predicted from the light-state PpSB1-LOV pdb coordinates (DSSP, [170]) is 2 % for α -helices, 5 % for β -sheets and 7 % for random coil and turn structures (Table 13).

Table 13: Comparison of the secondary structure content obtained from NMR data and the PpSB1-LOV X-ray structure

	secondary structure content		
	α -helix (%)	β -sheet (%)	random coil and others (%)
NMR (TALOS+) ^[114]	32.5	26.7	40.8
X-ray (DSSP) ^[170]	34.5	31.7	33.8
deviation (%)	2	5	7

Apart from the above described secondary structure estimation, TALOS+ provides a prediction for residue-wise backbone torsion Phi / Psi angles. Fig. 47 depicts the comparison of TALOS+ predicted backbone torsion angles (Phi, panel A and Psi, panel B) (shown in red) and the respective information obtained from the PpSB1-LOV X-ray structure (shown in black). With a few exceptions, the overall deviations between NMR predicted Phi, Psi torsion angles and the respective values obtained from the light-state crystal structure are small. Those exceptions include the region around the helix E α (containing the photoactive Cys53), the N-terminal end of the I β -strand, the H β -I β loop as well as other loop segments throughout the protein. Those deviations might well relate to the “mixed-state” nature of the PpSB1-LOV X-ray structure (see chapter 4.1.1). Interestingly, some of the residues showing the largest Phi / Psi angle deviations are localized in protein regions suggested previously to be involved in LOV photo-activation. Those residues include the photoactive Cys53 on E α as well as Q116 on I β which was suggested to be essential for the light-induced signal relay within LOV domains [171-174]. Details regarding the photo-activation mechanism of LOV domains will be discussed in chapter 4.4.

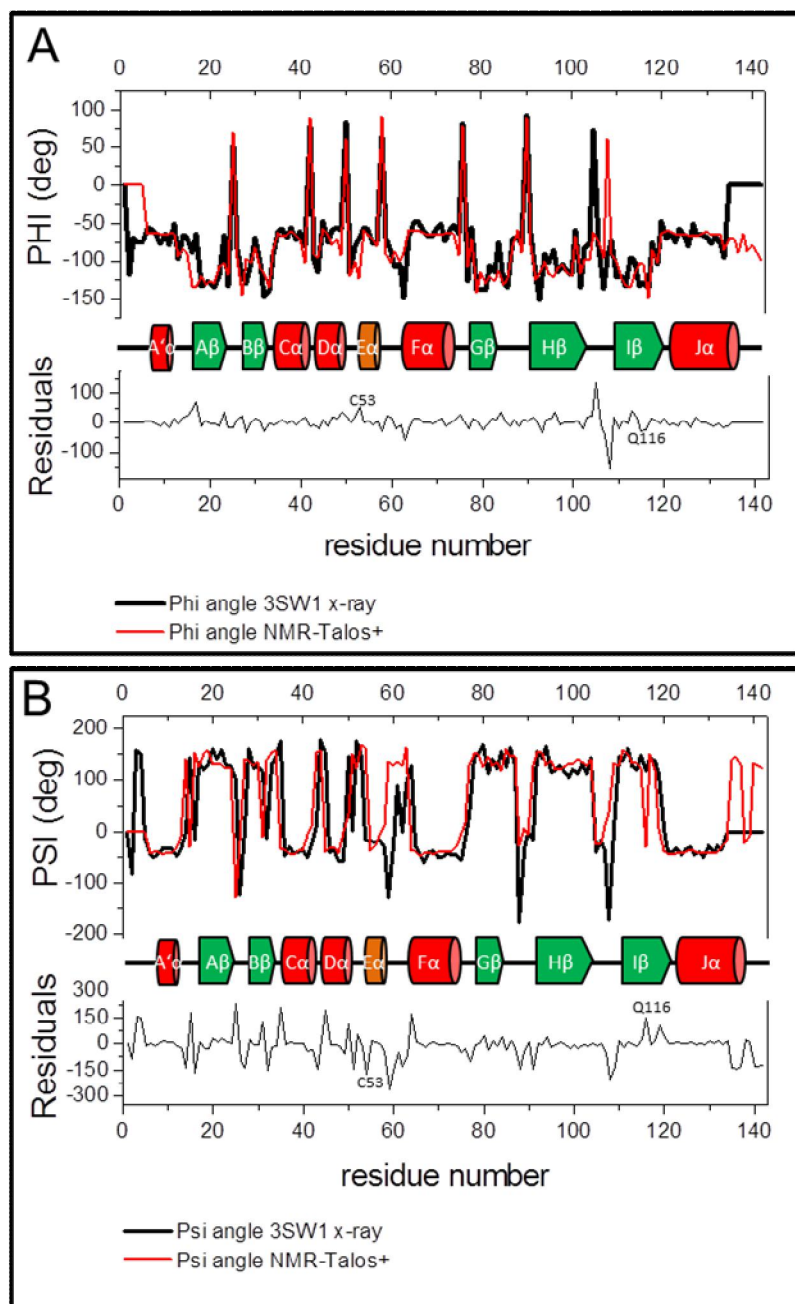


Figure 47: Comparison of TALOS+ derived Phi (A) and Psi (B) angles derived from PpSB1-LOV light-state NMR data (red line) and the PpSB1-LOV light-state X-ray structure (pdb entry 3SW1) (black line). The lower plot in each panel depicts the residue-wise difference (residuals) in Phi- and Psi angles between the NMR TALOS+ prediction and the values obtained from the respective crystal structure. Psi and Phi angles for the X-ray structure were calculated from the respective pdb coordinates using the YASARA structure suite [175].

4.2.2 H/D exchange and relaxation studies support the $J\alpha$ -helix orientation observed in the light-state X-ray structure of PpSB1-LOV

In different LOV systems, different $J\alpha$ -helix orientations have been observed. For AsLOV2 the $J\alpha$ -helix packs against the LOV core β -scaffold (“helix-in” conformation) shielding hydrophobic side-chains from solvent exposure in the monomeric protein [71]. In PpSB1-LOV the $J\alpha$ -helix protrudes from the core (“helix-out” conformation) and hydrophobic side-chains on the LOV β -scaffold are protected from the solvent by dimerization which is facilitated via hydrophobic interactions via the N-terminal cap ($A'\alpha$) and coiled-coil like interactions of the C-terminal $J\alpha$ -helix [157] (see also chapter 3.1 and 4.1).

The H/D exchange data presented in this study revealed that in both the dark- and light-state backbone amide protons of residues on the central β -scaffold are more protected from H/D exchange than those of residues throughout the rest of the protein (chapter 3.2.5.2). Those observations suggest shielding of the respective protein region from the solvent and would argue for a dimeric organization of the protein where solvent exposure is limited by subunit association as seen in the X-ray structure. In contrast, the backbone amide protons of residues located in the $J\alpha$ -helix are exchanged rapidly suggesting a “helix-out” orientation as seen in the light-state X-ray structure of PpSB1-LOV [157] and YtvA-LOV [73].

Moreover, relaxation data (T_1/T_2 and hetNOE) (3.2.5.1) as well as TALOS+ secondary structure predictions (see chapter 3.2.4 and 4.2.1) suggested that the $J\alpha$ -helix is well ordered in the light-state. At the same time relaxation data revealed a high mobility for the C-terminal Arg135-Ala142 amino acids (immediately after $J\alpha$) (3.2.5). The same was observed in the PpSB1-LOV X-ray structure where no electron density was found for the last few amino acids, indicative of high flexibility of the respective region. Also in the NMR structure obtained for the isolated PpSB1- $J\alpha$ peptide the same residues appear flexible (Fig. 19 in the results section 3.1.3.5).

Last but not least, as also discussed in chapter 4.1.2.1, SEC analyses suggested that PpSB1-LOV is dimeric in both the dark- and light-state in solution. In conclusion, the here presented light-state NMR data suggests a grossly similar overall structure and subunit

orientation as observed in the light-state X-ray structure of PpSB1-LOV. Some deviations are found in TALOS+ predicted Phi and Psi backbone torsion angles, i.e. for residues previously suggested to be involved in LOV domain photo-activation. Those differences might relate to the “mixed-state” nature of the PpSB1-LOV light-state X-ray structure (discussed in 4.1) However, high-resolution NMR derived structural models for the PpSB1-LOV protein in the light-state will be needed to unequivocally address this issue.

4.3 On the nature of PpSB1-LOV photo-activation

By comparing NMR data obtained for the light-state of the protein (see above) to the corresponding data recorded for the protein in the dark, information can be obtained about the structural consequences of PpSB1-LOV photo-activation. Therefore, dark-state 3D HNCO, as well as relaxation and H/D exchange experiments were measured and compared for both states of the protein.

4.3.1 Chemical shift perturbation analyses suggest a potential mode for LOV-photo-activation

Comparison of spectral quality revealed, that in the light-state the protein seems to be more susceptible to proteolysis and or degradation than in the dark-state. After 5 days illumination, evidence for degradation or conformational heterogeneity, i.e. appearance of additional not assignable cross peaks, were seen in the respective HSQC spectra (Fig. 40, chapter 3.2.3.2). In contrast, in the dark the protein seems stable. The corresponding HSQC spectra showed only minor changes after several days of measuring time at 40 °C (Fig. 39, chapter 3.2.3.1). Similar effects have been observed in limited proteolysis experiments performed with the oat phototropin1 LOV2 domain (AsLOV2) [71, 176]. Nevertheless, the PpSB1-LOV protein is stable enough to allow a complete backbone resonance assignment (99 % complete for both states) and an almost complete side-chain proton assignment for the light-state (89 %). The assignment of both states allowed us to identify regions or residues of the molecule that experience light-dependent changes. In order to quantify those changes we carried out chemical shift perturbation analyses (Fig. 48 and Fig. 49). Comparisons of ^1H - ^{15}N HSQC spectra showed that illumination produces widespread changes in backbone chemical shifts (Fig. 48). For maximum resolution, dark- and light-state 3D HNCO spectra were compared, using the minimum chemical-shift difference method [144] to quantify the overall structural

changes. The combined minimum chemical shift perturbation ($\Delta\delta_{\min}$) for the backbone nuclei (NH, C, CO) was calculated as described in Materials and Methods (2.11.7) and the results are plotted in Fig. 49 (panel A).

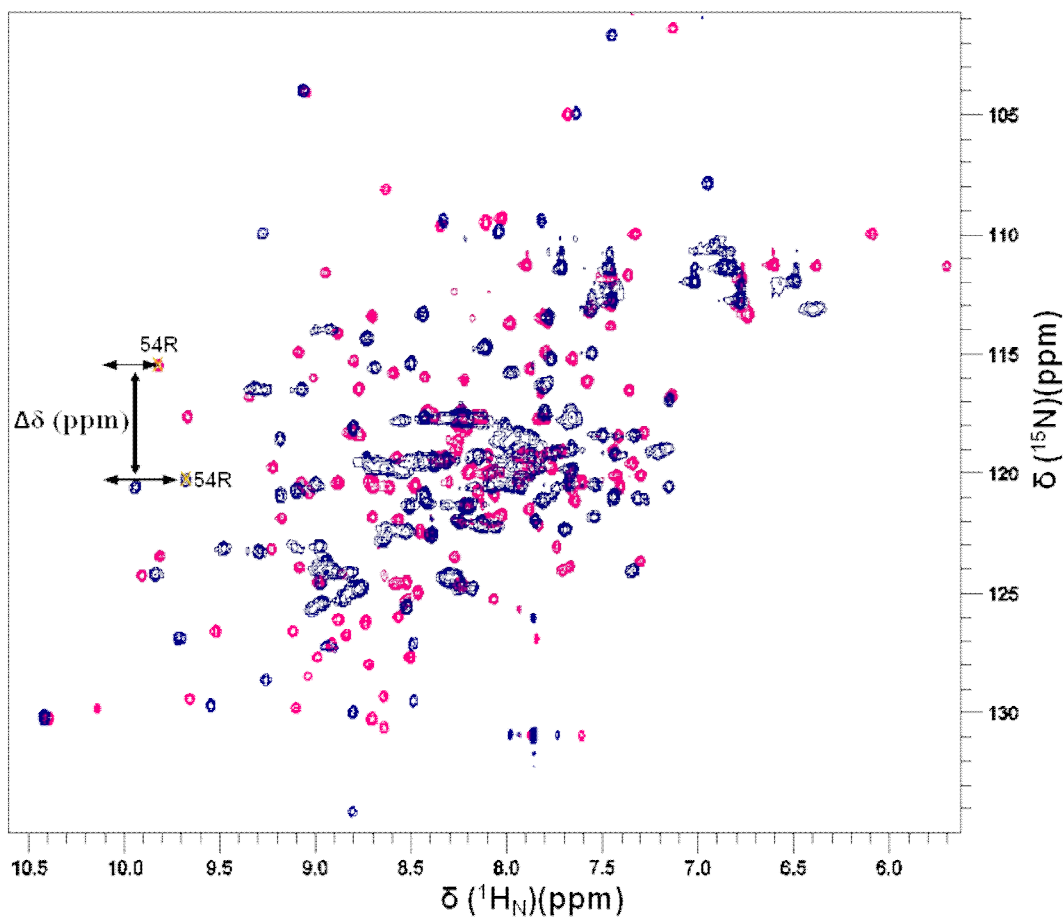


Figure 48: Schematic illustration of the NH chemical shift change for residue 54R (vertical black arrow) of PpSB1-LOV illustrating the peak “movement” between light- (pink) and dark state (blue).

Subsequently, those $\Delta\delta_{\min}$ values were mapped onto the light-state crystal structure of PpSB1-LOV (Fig. 49, panel B and C).

It is worth noting that chemical shift changes are consistently observed throughout most of the primary sequence of the protein. In the LOV core, the largest differences were observed for residues in the immediate vicinity of the FMN chromophore. FMN-proximate regions experiencing strong chemical shift changes include the E α -helix (D52 - Q57), which contains the photoactive Cys53, and residues on the central β -scaffold (A β , B β , H β and I β). Strong chemical shift changes are also observed for residues located more distal to the FMN chromophore. Those regions include residues on the D α -helix, residues on A β and B β , as

well as residues on the N-terminal A' α capping helix. Smaller, but well detected, changes are observed for the first few residues of the protruding J α -helix.

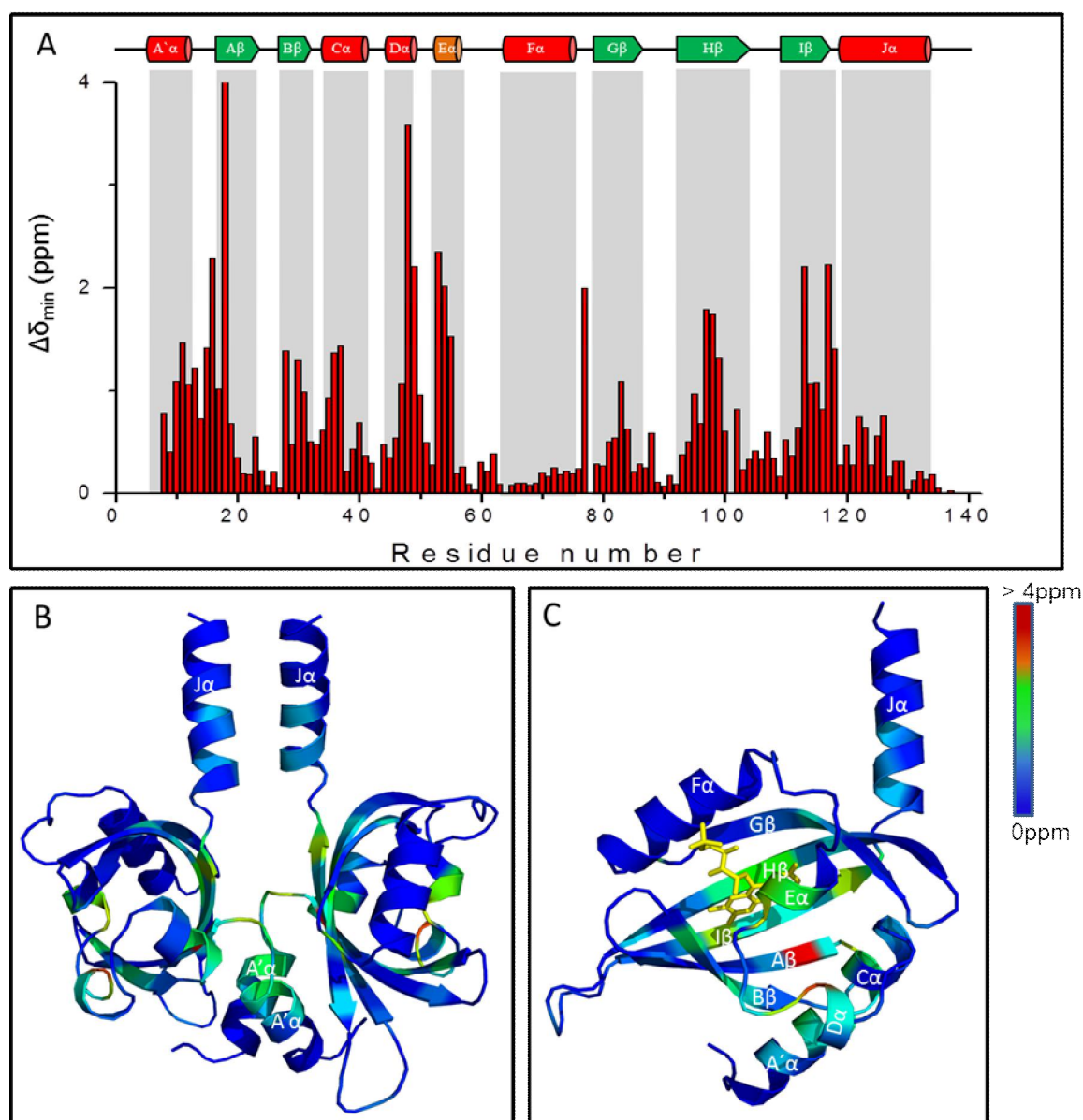


Figure 49: (A) Minimum chemical shift perturbation ($\Delta\delta_{\min}$) obtained from the 3D HNCOC spectra recorded for the dark- and light state of PpSB1-LOV. $\Delta\delta_{\min}$ values were calculated according to $\Delta\delta_{\min} = [\Delta\delta(^1\text{H})^2 + 0.17\Delta\delta(^{15}\text{N})^2 + 0.39\Delta\delta(^{13}\text{C})^2]^{1/2}$. The characteristic secondary structure elements are shown in the upper part of the panel. The unassigned N-terminal region is not included in the plot. (B) Minimum chemical shift differences were mapped onto the light-state crystal structure of PpSB1-LOV (pdb id: 3SW1). (C) One chain of (B) reoriented for clarity. FMN is represented in yellow sticks. On upper left corner the color gradient encoding the magnitude of the shift change is shown.

Chemical shift changes are more or less absent in the loop region connecting the helices E α and F α as well as on the complete F α -helix. This region is of particular interest as it contains

two arginine residues (R61 and R66) which are involved in the tuning of the PpSB1-LOV dark recovery reaction. Substitution of R61 and R66 for the respective amino acids present in the fast reverting PpSB2-LOV protein (H61, I66) accelerated the dark recovery reaction by about 280-fold compared to the wild-type protein [49, 157]. Moreover, in the light-state X-ray structure the E α -F α loop shows an unusual “bent-in” conformation caused by the directional pull of the FMN phosphate on the R61 side-chain. This results in the formation of a salt-bridge between the N ϵ atom of R61 and the terminal FMN phosphate [157]. Absence of strong chemical shift changes for residues on this loop thus indicates that this “bent-in” conformation observed in the light-state X-ray structure is not caused by photo-activation, but seems to be an intricate feature of PpSB1-LOV in both states.

Taken together, light-induced conformational changes are present throughout most of the protein but appear concentrated on the β -scaffold (i.e on A β , B β , H β - and I β). This highlights the importance of the respective region for PpSB1-LOV photo-activation. For other LOV systems such as the light-dependent DNA-binding protein EL222 as well as for AsLOV2 the central β -scaffold was likewise reported to suffer from light-induced conformational changes [71, 177]. Interestingly, the two auxiliary structural segments A' α and J α appear to undergo a light-correlated conformational change.

Thus it is tempting to speculate that intra-molecular signal-relay might occur from the immediate vicinity of the light absorbing FMN chromophore (β -scaffold) (where incident photons are captured to initiate FMN-Cys53 adduct formation) to the subunit interface of the dimer (A' α) and consequently also to solvent exposed surface regions such as the J α -helix.

In the following we analyzed the PpSB1-LOV light-state X-ray structure with regard to a possible mode of intra-molecular signal relay from the FMN chromophore to more distal regions such as A' α and J α . Hereby, the NMR derived light-induced chemical shift changes are considered as a marker for a structural change of the respective residue, although our NMR data in fact only provides information about the backbone nuclei and cannot account for potential side-chain movements. Moreover, the obtained data is discussed in light of generally accepted mechanisms which were suggested for LOV domain photo-activation.

Apart from FMN-C53 adduct formation, illumination of LOV domains is suggested to induce a rearrangement of the side-chain of a highly conserved glutamine residue (Q123 in YtvA [171], Q513 in AsLOV2 [172], Q116 in PpSB1-LOV). In detail, based on FT-IR, NMR and UV/Vis data, it was proposed that in the dark the respective glutamine forms a hydrogen-bond to the FMN-O4 atom while upon illumination this hydrogen-bond is broken, the side-chain flipped and a new one established with the newly protonated FMN-N5 atom [73, 100, 101]. For AsLOV2 it was argued that alteration of the FMN-Q513 hydrogen bonding mode could influence the structure of the anchoring I β strand, which in turn induces displacement and unfolding of the J α -helix which in the dark packs against the β -scaffold (i.e against I β) [172].

Fig. 50 depicts the immediate vicinity of the PpSB1-LOV FMN chromophore and highlights residues experiencing strong chemical shift changes (color coded as in Fig. 49). Hydrogen-bonding contacts between the FMN chromophore and the surrounding amino acids are shown, as are hydrogen-bonds that connect structural elements displaying strong chemical shift changes.

Chemical shift perturbation analyses reveal strong chemical shift changes for the backbone nuclei of Q116 and N95 both of which are hydrogen-bonded to the FMN ring-system (Fig. 50). Correlated changes are detected for neighboring charged residues such as K117 and E96 on the respective β -strands (I β and H β). Moreover, hydrophobic residues on A β (I18, V19) and I β (I113) appear perturbed by photo-activation. The side-chains of all of the latter residues point outwards into the subunit interface in the region of the A' α -helix and the beginning of the J α -helix. For orientation the distance between the C53-S γ atom and the FMN-C4a atom as well as the FMN chromophore (in line representation) are shown in Fig. 50.

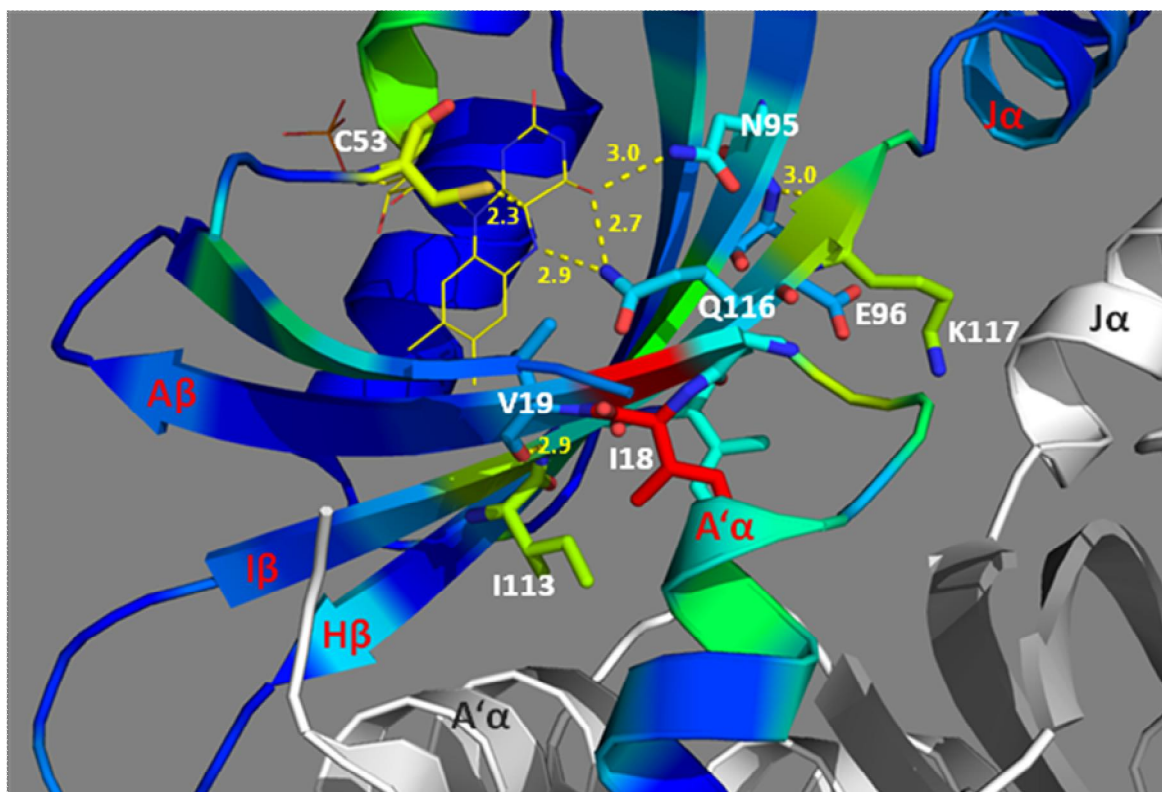


Figure 50: Vicinity of the FMN chromophore of PpSB1-LOV. The light-state X-ray structure with bound FMN chromophore is shown. The color-code is the same as is Fig. 49 and visualizes chemical shift changes ($\Delta\delta_{\min}$) obtained from chemical shift perturbation analyses. The FMN chromophore is in shown in yellow lines. For orientation the distance between the C53-Sy atom and the FMN-C4a atom (2.3 Å) as well as the second PpSB1-LOV subunit (in grey) is depicted. Polar residues that interact with the chromophore (Q116, N95) are shown. Hereby, Q116 on strand I β forms hydrogen bonds with the N5 atom and the O4 atom of the FMN isoalloxazine ring. N95 (on H β) is hydrogen-bonded to O4. All residues in the vicinity of the FMN chromophore that experience strong chemical shift changes are shown in stick representation colored according to the magnitude of the chemical shift change. Hydrogen-bonds (yellow broken lines) that connect the FMN chromophore to surrounding residues as well as H-bonds connecting the respective secondary structure elements are shown.

Given those observations a hypothesis regarding the intra-molecular signal relay can be presented:

- i) Light absorption in the FMN chromophore causes covalent bond formation between the FMN-C4a and the C53-Sy atom.
- ii) This in turn might results in a rearrangement of the hydrogen-bonding network surrounding the chromophore, potentially involving Q116 and N95. The conformational change, that those two residue experience might in turn influence the structure of their

anchoring β -strands (I β and H β), which is detected as strong chemical shift change at the respective backbone nuclei (Fig. 50).

iii) Moreover, A β , H β and I β are interconnected via H-bonds over the backbone resulting in correlated chemical shift changes in all three protein regions.

iv) In all three segments strong chemical shift changes are seen for the backbone nuclei of polar (E96, K117) and hydrophobic residues (I18, V19 and I113) whose side-chains point outwards into the PpSB1-LOV subunit interface i.e contacting distal protein regions such as the A' α and the J α -helix.

v) Those distal surface- (A' α) or solvent-exposed (J α) structural elements might be rearranged e.g. to alter the surface properties (hydrophobicity, electrostatic potential) of PpSB1-LOV which in turn could affect interaction affinities with potential, yet to identify, downstream signaling partners.

4.3.2 Further NMR spectroscopic evidence for the proposed intramolecular signal-relay mechanism

In the previous chapter the PpSB1-LOV β -scaffold i.e. the strands A β , H β and I β were proposed as key-transducer elements facilitating both LOV photo-activation and intramolecular signal relay. In the following paragraphs additional evidence will be presented for the involvement of those structural regions in photo-activation.

4.3.2.1 NMR-based secondary structure predictions for the dark- and light-state of PpSB1-LOV suggest local secondary structural changes in A β -B β region

TALOS+ secondary structure analyses were carried out for the dark-state of PpSB1-LOV. Overall the TALOS+ predicted secondary structure for the dark-state of PpSB1-LOV is very similar to the above presented light-state data (see chapter 3.2.4 for a direct comparison). Nevertheless, deviations are identified for some of the above described key structural regions.

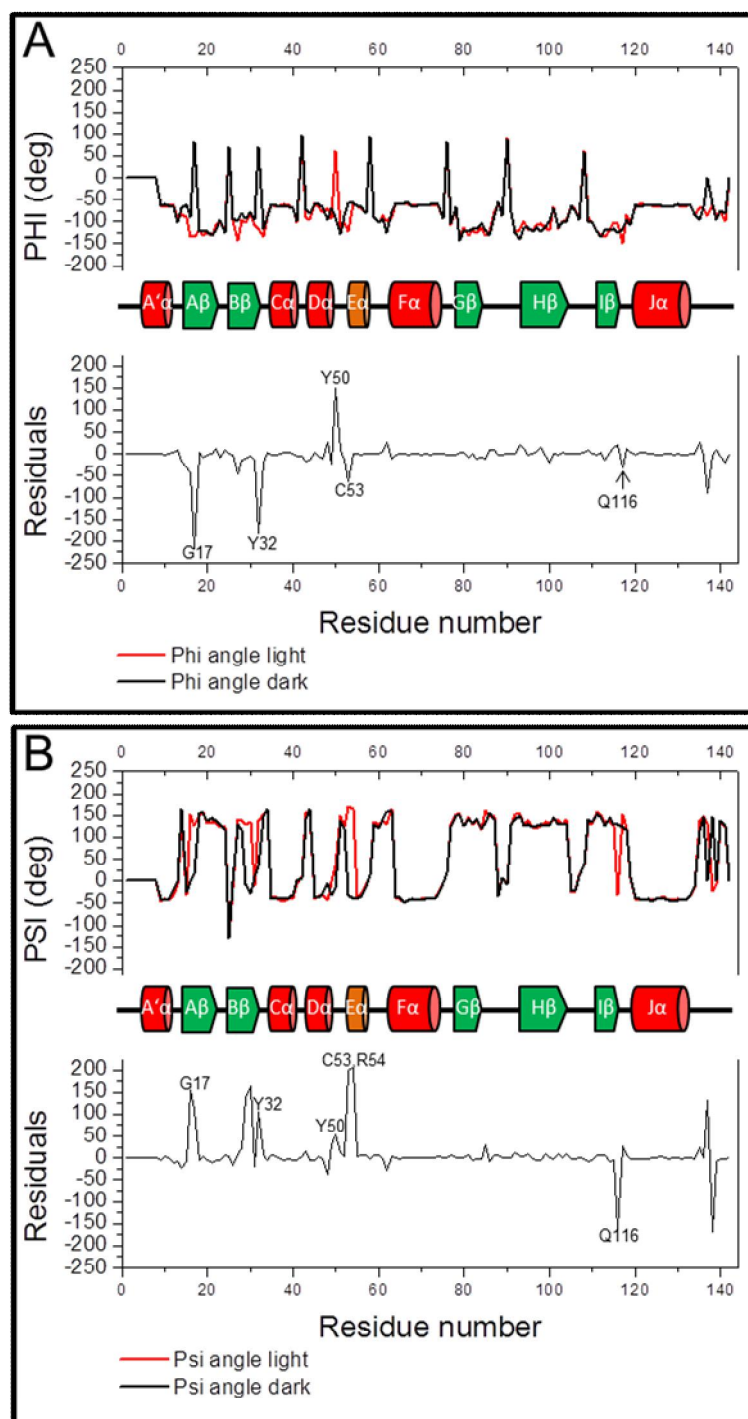


Figure 51: Comparison of TALOS+ derived dark- (black line) and light-state (red-line) Phi (A) and Psi (B) backbone torsion angles derived from NMR data. The lower plot in each panel depicts the residue-wise difference (residuals) between the dark- and light-state Phi- and Psi angle predictions.

In Fig. 51 Phi (panel A) and Psi (panel B) backbone torsion angles predicted from the NMR data of the dark (black line) and light-state (red line) are compared. The difference (light-

dark) is shown as residuals below the respective plot. Major differences are seen for the residues G17 (on A β), Y32 (B β), Y50, C53 and R54 (on E α) as well as for Q116 (on I β).

The most pronounced difference between dark- and light-state secondary structure is found for the segment comprising the B β strand. Here, TALOS+ predicts an unstructured region extending from residue I29 to residue Y33 which is predicted to form a β -strand (B β) in the light state.

Apart from providing secondary structure predictions for NMR data, TALOS+ also provides measures as to how reliable those predictions are. For example the software provides for each residue an estimation about the reliability of the secondary structure assignment (unambiguous/good, ambiguous, dynamic, bad). Those estimations are based on a set of strict rules. In brief, a prediction is judged as “good/unambiguous” when [114]:

- 1.) All 10 best database matches fall in a consistent region of the Ramachandran plot.
- 2.) The confidence of the artificial neural network (ANN) 3-state Phi/Psi distribution prediction for a given residue (defined as the difference between the probabilities of the two most favored predicted states) must be above 0.6. (0.7 for residues with "Positive-Phi" prediction).
- 3.) The RCI-predicted order parameter S² value should be above 0.5. All the cases with predicted S² value <0.5 are likely to be "Dynamic", and will not be considered as unambiguous predictions.

Given those rules, TALOS+ judges the prediction for the B β region as reliable (“good/unambiguous”) for both the dark- and light-state NMR data of PpSB1-LOV. However, please bear in mind that over a database of 200 proteins, about 2.5 % of unambiguous TALOS+ predictions were incorrect relative to the corresponding crystal structure [114]. Hence, erroneous TALOS+ predictions (even when the prediction is judged as “good”) cannot be ruled out completely.

Therefore, in order to validate the observed loss of secondary structure for the B β -strand in the dark-state several additional analyses were carried out.

The chemical shifts of the H α protons in NMR experiments can be used to obtain information about the secondary structure of proteins, since these values are extremely sensitive to the chemical environment of the proton. It has been shown that the sign and magnitude of the proton chemical shifts compared to the corresponding random coil values are correlated to the secondary structure element in which the respective residues resides [178]. For example, H α 's that are part of an α -helix display considerably upfield shifted chemical shift values compared those tabulated for random-coil peptides [138]. Correspondingly, residues belonging to a β -sheet display downfield shifted chemical shift values. Based on this, Wishart *et al.* (1992) developed the chemical shift index (CSI) method for secondary structure prediction from NMR data [178]. Fig. 52 depicts the comparison of TALOS+ (panel A, B) and CSI (panel D, E) secondary structure predictions for the light (panel A, D) and dark-state (B, E) of PpSB1-LOV. In CSI predictions β -strand structure is defined when three or more "+1" HA and/or "-1" CA/CO relations are sequentially found. According to this observation the CSI prediction also defines the region from residue I29 to residue Y33 as unstructured in the dark state and suggests formation of a β -strand (B β) upon illumination.

Ring-current shifts are one of the main reasons to cause chemical shift dispersion of ^1H -NMR signals of proteins in their folded states. Hence the effect of the ring-current of neighboring aromatic residues on the chemical shift values recorded for the B β region has to be considered as a cause for the apparent loss of secondary structure of the B β region in the dark-state, i.e. by shifting the respective resonances to unexpected values. Fig. 55 (panel C) shows the ring-current shifts of the C α protons (red line) together with the amide proton ring-current shift (black line) calculated based on the light-state crystal structure of PpSB1-LOV using the program MOLMOL [145]. According to MOLMOL analyses the ring-current shift for the respective residues (I29-Y33) is minimal (Fig. 52, C). However, without a dark-state crystal structure or high-resolution dark-state NMR models of PpSB1-LOV we cannot rule out that ring-current shifts for the dark-state substantially deviate from the ones predicted here from the light-state structure.

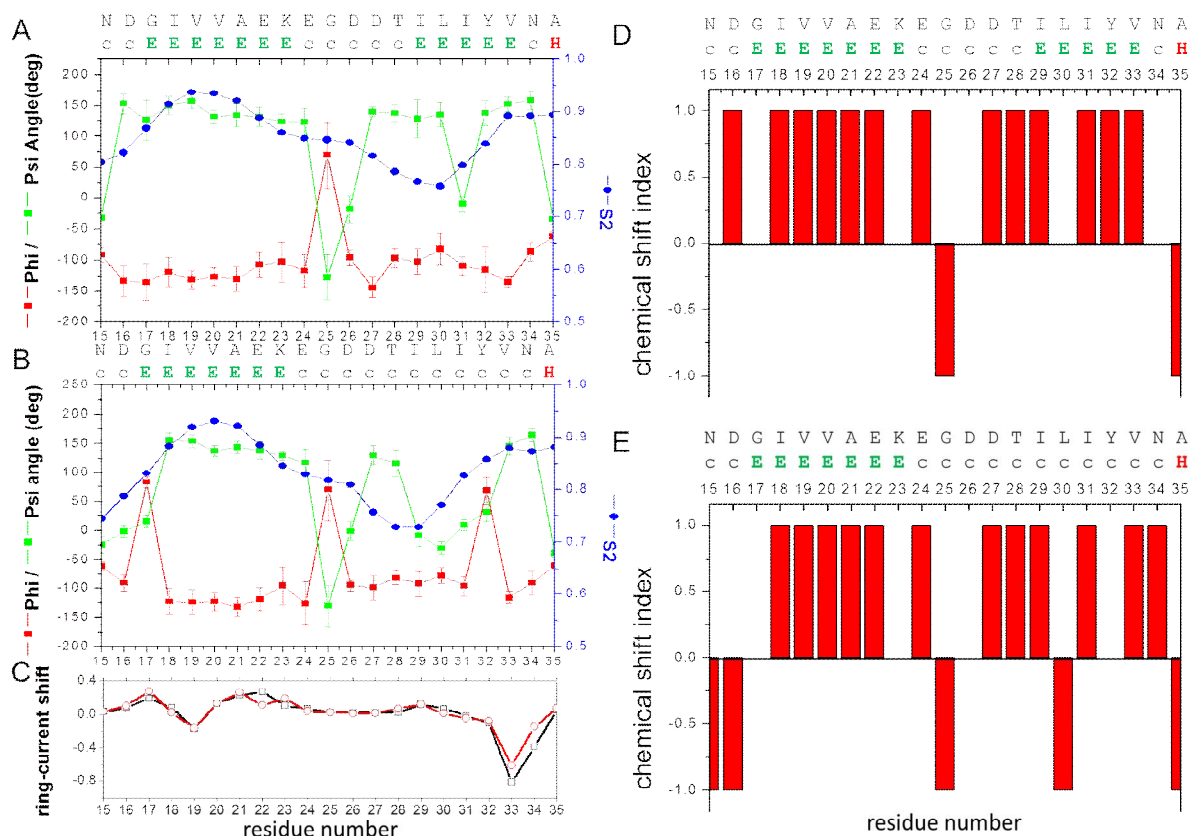


Figure 52: Summary of TALOS+ secondary structure predictions for the Aβ-Bβ region (N15-A35) of PpSB1-LOV in the light (A) and dark-state (B). Backbone torsion angles were calculated using TALOS+. Phi and psi angles are shown in red and green, respectively. Additionally, the order parameter S_2 predicted using the random coil index (RCI) approach [179] is depicted (blue). (C) MOLMOL predicted ring-current shift values for the C α (red) and amide protons (black) Panel (D) and (E) depict the consensus CSI graph obtained for the dark- and light-state NMR data, respectively. Three or more +1 CSI values are indicative for β -strand structures.

Interestingly, when comparing the light-state PpSB1-LOV X-ray structure to dark-state X-ray structures of other LOV proteins such as e.g. *B. subtilis* YtvA and *C. reinhardtii* phototropin LOV1, structural differences in the A α -B β loop region and consequently the B β -strand are apparent. Compared to YtvA in the dark-state, an extended anti-parallel arrangement of the A β and B β -strands as well as a concomitant shortening of the A β -B β loop can be seen in the PpSB1-LOV light-state X-ray structure (Fig. 53).

Thus the presented NMR data, i.e. suggesting the loss of secondary structure for the B β -strand in the dark-state of PpSB1-LOV, might suggest that this structural region plays an important role in PpSB1-LOV photo-activation. Consequently, in the dark, the region of the

β -strand might be more flexible and or unstructured, whereupon illumination a β -strand conformation, as seen in the X-ray structure, might become established or fully folded.

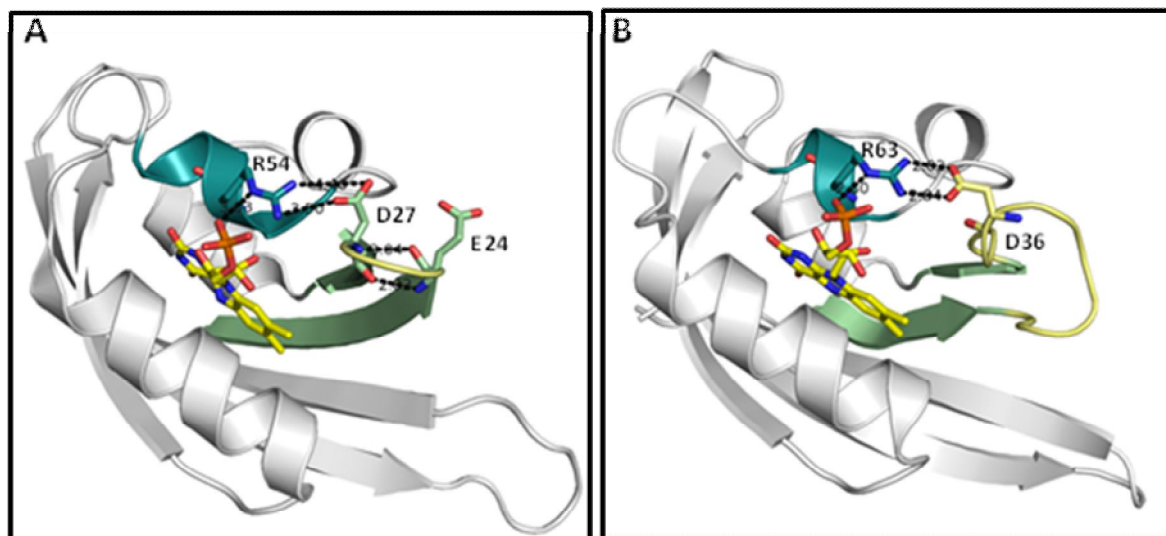


Figure 53: Comparison of the A β -B β arrangement in (A) the light-state of PpSB1-LOV (pdb entry: 3SW1) and (B) the dark-state of YtvA-LOV (pdb entry: 2PR5). Conserved residues (D27 in PpSB1-LOV, D36 in YtvA-LOV) as well as the conserved FMN-phosphate binding arginine (R54 in PpSB1-LOV, R63 in YtvA-LOV) are shown in stick representation. Additionally in both PpSB1-LOV and YtvA-LOV the latter arginine forms H-bonds to D27 or D36, respectively. The D27-E24 backbone H-bonding interaction that facilitates the A β -B β arrangement as seen in PpSB1-LOV is additionally shown in panel (A).

In conclusion, comparative TALOS+ analyses of dark- and light-state NMR data corroborate the functional importance of the PpSB1-LOV β -scaffold for photo-activation and or signal-transduction.

4.3.2.2 Relaxation experiments

Overall, the difference in the relaxation behaviour of PpSB1-LOV in the dark- and light-states is small. In both states the LOV core domain as well as the N-terminal A' α helix are quite rigid, whereas the last 8 amino acids, immediately after the J α -helix, appear flexible (high R1/R2 values, low or negative hetNOEs).

From the respective dark- and light-state (R1/R2) relaxation data (see chapter 3.2.5.1), light-dark difference R1/R2 values are calculated and mapped onto the PpSB1-LOV light-state X-ray structure (Fig. 54). Overlapping and unassigned peaks are deleted from both light- and the dark-state data. Small (light-dark) differences in (R1/R2) values can be found for regions previously implied in LOV photo-activation (see chapter 4.3.1 and 4.3.2). Those regions include the A' α -helix and the immediate end of the I β -strand. Additionally, the most

pronounced difference in dark vs. light relaxation behavior is found for residues within the H β -I β loop (see also chapter 3.2.5.1). In previous X-ray crystallographic studies this loop segment was shown to adopt variable orientations in different proteins. Hence it seems likely that the H β -I β loop is intrinsically more flexible than the rest of the LOV core domain.

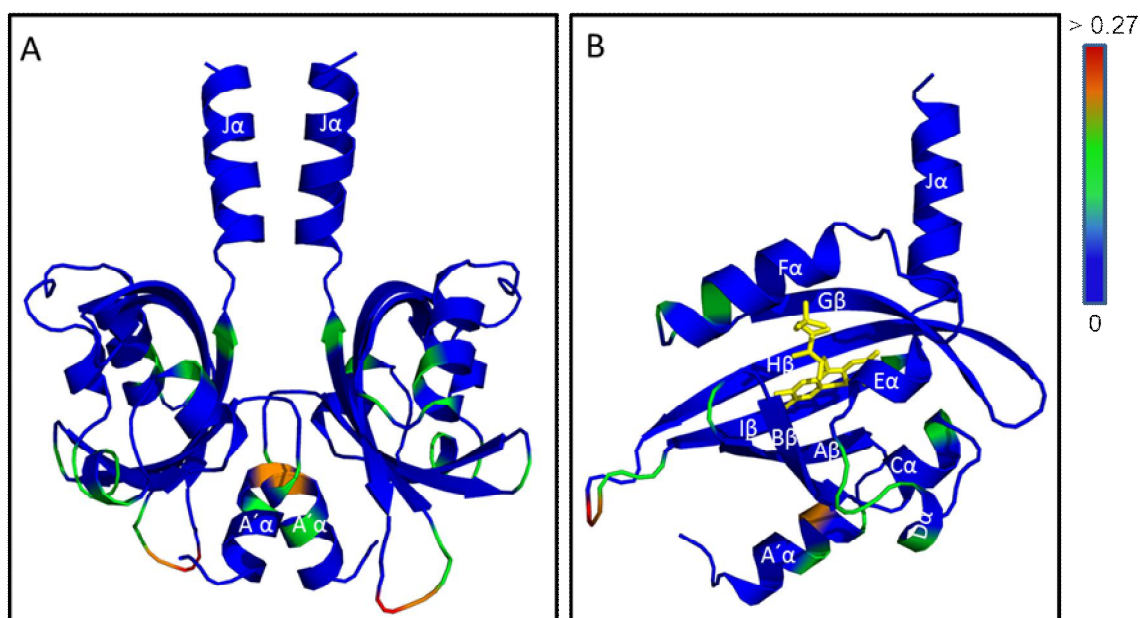


Figure 54: Light-dark relaxation data (R1/R2) mapped onto the light-state PpSB1-LOV X-ray structure (A) In panel (B) one chain of (A) is shown, reoriented to provide a better view into the chromophore binding pocket. The FMN chromophore is shown in stick representation (yellow). In upper left corner the color gradient encoding the magnitude of the (R1/R2) change is shown.

4.4 Implications of the here presented NMR data for the conservation of the LOV photo-activation and signal-relay paradigm

The here presented dark- and light-state NMR data provides for the first time a detailed view into potential PpSB1-LOV photo-activation and signal-relay mechanisms. Although the here presented data does not allow unequivocal conclusions as to the exact mechanism of PpSB1-LOV photo-activation and intra-molecular signal-relay, several hypothesis were brought forward in the previous chapters. Given those observations immediate questions arise regarding the conservation of the here proposed mechanisms with respect to other so far structurally characterized LOV photosensory systems. Thus, in order to better understand the molecular basis of PpSB1-LOV signaling and to assess whether LOV proteins from different species employ a common signaling mechanism, older and more recent hypothesis regarding

LOV photo-activation and signal-transduction will be re-reviewed in the following chapters. In the concluding paragraph this information will be placed in the context of PpSB1-LOV photo-activation, intra-molecular signal relay and potential downstream signal-transduction. In Fig. 55 a set of so far proposed signal-transduction mechanisms are schematically summarized.

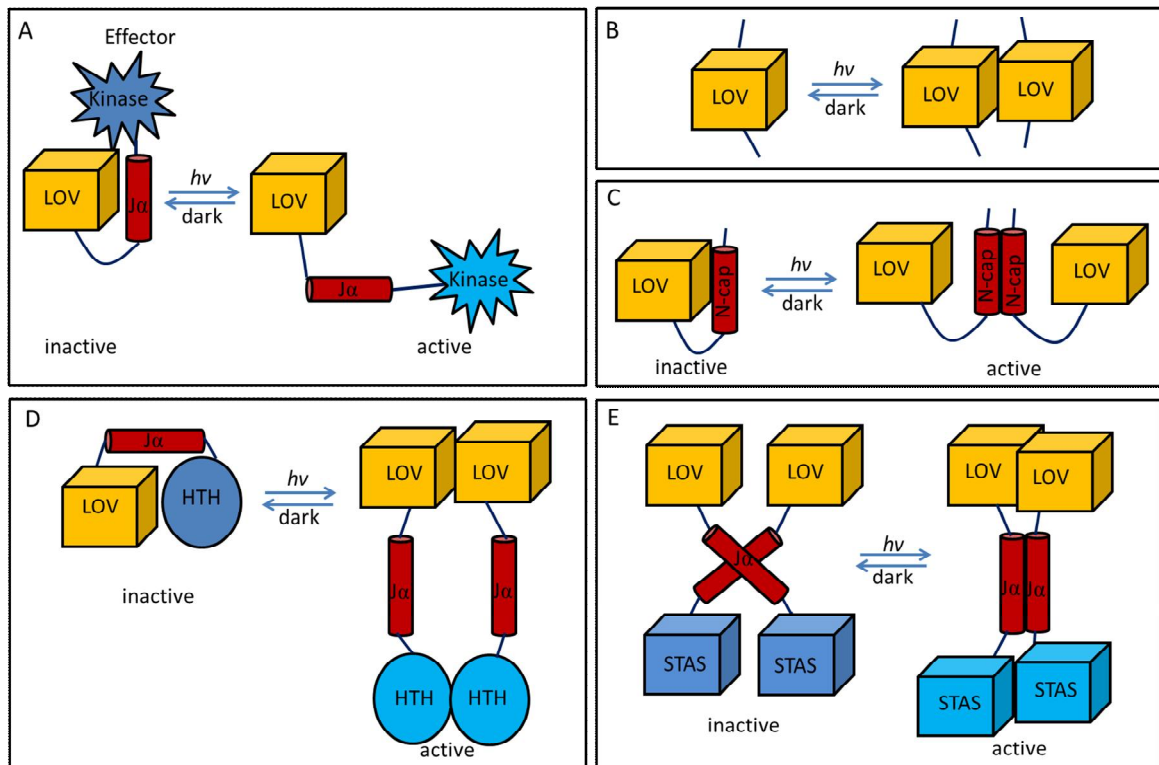


Figure 55: Proposed signal transduction mechanisms for pro- and eukaryotic LOV proteins. **(A)** In plant phototropins, photo-activation induces a conformational change in the LOV2 domain that results in disruption of the interaction between the LOV core and the J α -helix. This leads to activation of the effector domain [71]. **(B)** The second mechanism proposed for phototropin activation involves a transient light-induced dimerization of the full-length protein via interaction of the LOV1 and LOV2 domains [128]. **(C)** In the VVD protein of *N. crassa*, photo-activation leads to a rearrangement of the N-terminal cap followed by dimerization of the protein [173]. **(D)** For the LOV-HTH protein EL222 of *E. litoralis*, it has been proposed that photo-activation disrupts the interaction surface between the LOV domain and the HTH domain. Illumination hereby weakens the contact, thus promoting dimerization of the freed HTH domains which in turn facilitates target DNA binding [177]. **(E)** In the YtvA protein from *B. subtilis*, photo-activation was suggested to induce a minor rotation of the two LOV subunits relative to each other which results in J α helix and consequently effector (STAS) domain reorientation [180].

4.4.1 Photo-activation of the plant phototropin LOV2 domain involves dissociation of the J α -linker helix

As briefly outlined in the introduction (chapter 1.2.2.3), plant phototropins are multi-domain photosensory systems, consisting of two N-terminal LOV domains (LOV1 and LOV2) associated with a C-terminally fused Ser/Thr kinase. Upon illumination, the phototropin kinase is autophosphorylated. Signal-transduction in the cell was suggested to occur via a, yet poorly characterized, phospho-relay cascade that initiates blue-light dependent plant responses such as phototropism and chloroplast relocation [181, 182].

Initial X-ray crystallographic studies on the isolated LOV1 and LOV2 domains of plant and algal phototropins were little conclusive, as photo-activation of dark grown crystals of the *Adiantum capillus-veneris* Neochrome (Phy3) LOV2 domain [66] and the *C. reinhardtii* phot LOV1 domain [101] did not reveal large scale conformational changes upon illumination. Photo-activation and intra-molecular signal-relay in full-length phototropin1 was first studied in detail by solution NMR techniques using an extended LOV2 domain construct obtained from the *A. sativa* phototropin1 molecule [76]. Those studies, by Harper *et al.* in (2003), suggested that upon photo-activation of the LOV2 domain a C-terminally located auxiliary structural element, namely the J α -helix dissociates from the LOV core and becomes unfolded [71]. The authors suggested, and later proved by mutagenesis [74], that J α -helix unfolding eventually results in kinase activation in full-length phototropin, which in turn enables downstream signaling responses and thus plant phototropism [71, 74] (Fig. 55, panel A). Another mechanism brought forward to achieve light-dependent regulation of phototropin kinase activity was suggested to involve a light-dependent change in the quaternary structure of the *Arabidopsis* phototropin1 LOV2 domain. The authors suggested that photo-activation of LOV2 causes transient dimerization of the isolated sensory LOV domain, which would trigger full-length phototropin dimerization and thus result in kinase activation and phototropism [128] (Fig. 55, panel B). Whether the latter proposed mechanism involves light-dependent J α -linker helix unfolding has never been studied and remains to be addressed. Recent NMR work suggested that the A' α -helix present N-terminally to the LOV2 core domain, is equally important for LOV2 photo-activation [76]. The authors suggested that a relay mechanism between the N-terminal A' α -helix and the C-terminal J α -helix exists which is important for J α dissociation [76].

4.4.2 Light-induced transient dimerization of the short LOV protein VVD triggers protein-protein interactions in *N. crassa*

The short LOV protein VVD from *N. crassa* consists of a single LOV domain and an N-terminal extension, termed N-cap. The N-cap contains a short helix similar to the A' α -helix of PpSB1-LOV. Like PpSB1-LOV, VVD does not contain a fused effector domain [54].

X-ray crystallographic studies, as well as solution experiments using SEC and Small-Angle X-ray Scattering (SAXS) approaches, suggested that VVD is monomeric in the dark, with the N-cap packing against the LOV-core β -scaffold [54]. Illumination triggers disruption of the LOV-core / N-cap interaction [54] and results in the transient formation of VVD dimers [173]. This enables the light-dependent interaction between VVD and the Whitecollar complex (WCC) of *N. crassa* (Fig. 55, panel C). Although, the conformational change in the N-cap region is important for the function of VVD [54], the exact mechanism as to how the light-dependent dissociation of the VVD N-cap and the subsequent VVD dimerization, triggers WCC interaction is unknown.

4.4.3 Light-induced DNA binding by an LOV-HTH protein is mediated by sensor / effector domain rearrangement

The recently identified EL222 protein from *E. litoralis* represents a light-regulated DNA-binding protein. It consists of an N-terminal LOV domain flanked by an N-cap segment (including A' α -helix) and a C-terminal $J\alpha$ -linker helix. The latter one linking the sensor LOV domain to a helix-turn-helix (HTH) DNA-binding domain [177]. The dark-state crystal structure of full-length EL222 revealed a monomeric protein, where the HTH domain is docked against the β -scaffold of the LOV domain [177]. Solution NMR techniques were used to probe EL222 photo-activation and intra-molecular signal-relay [177]. Apart from changes in the immediate vicinity of the FMN chromophore of the protein, chemical shift perturbation analyses revealed structural changes as far as >15 Å away from the chromophore. Those changes were localized in the A' α -helix as well as in the HTH domain itself. Both latter structural segments are localized in close proximity to the LOV β -scaffold of EL222. Most importantly, no unfolding of the $J\alpha$ -helix as seen in AsLOV2 (see above 4.4.1) was observed for EL222. Based on a comparison to other HTH DNA-binding proteins, it was suggested that EL222 should bind to its DNA target as a dimer (Fig. 55, panel D). The authors thus

suggested that photo-activation of the EL222 LOV domain results in disruption of the interaction between the LOV and HTH domains, which frees the HTH domain and results in dimerization of EL222 and DNA binding [177].

4.4.4 Light-induced rotation of coiled-coil forming J α -linker helices in the dimeric YtvA photoreceptor from *B. subtilis*

The *B. subtilis* YtvA photoreceptor consists of an N-terminal LOV domain fused to a C-terminal sulfate transporter anti-sigma-factor antagonist (STAS) domain. In the dimeric protein, the LOV core is preceded by a short A' α -helix (N-cap) [180] and LOV and STAS domains are connected via a J α -linker helix [73, 180]. YtvA and related proteins play an important role in the blue-light-mediated stress response of *B. subtilis* and other Gram-positive bacteria such as *Listeria monocytogenes* [183].

A few years ago the crystal structure of the YtvA-LOV domain comprising the entire J α -helix but lacking the N-cap region was solved in both the dark- and light-state. As for most LOV X-ray studies, crystals were grown in the dark and photo-activation was achieved by illuminating the crystals before diffraction data was collected. In both crystal and solution, the isolated YtvA-LOV domain as well as the full-length protein is reported to be dimeric [73, 180, 184, 185]. In the crystal YtvA-LOV forms a parallel dimer with protruding J α -helices, which however do not form a distinct coiled-coil like motif as seen for the structurally similar PpSB1-LOV protein [157]. Although, not directly observed in the crystal, it was later proposed that the YtvA J α -helices can adopt coiled coils in the dimeric full-length photoreceptor. This notion was based on bioinformatics analyses and a protein-engineering studies, were structurally similar PAS domains of the oxygen-sensor FixL were replaced by the YtvA-LOV domain [186]. Using this approach an artificial light-gated kinase could be generated whose functionality was controlled by light [186]. By varying the J α -linker composition of those constructs, the authors observed a heptad repeat pattern for the functional response, i.e. the possibility for light activation of the FixL kinase by the LOV domain [186].

Based on the data, obtained from illuminated crystals of YtvA-LOV it was moreover suggested that photo-activation causes the two LOV domain monomers to rotate relative to each other resulting in a quaternary structural change that can be described as a scissor-like

movement [73]. In the full-length photoreceptor this minor (4-5 °) rotation would be transmitted via the rigid J α coiled coil and translated into a displacement of the YtvA-STAS domains relative to each other in the dimer [180]. This in turn might allow for altered protein-protein interactions to occur in the cell upon photo-activation and hence enable the observed physiological response [180] (Fig. 55, panel E).

4.5 Conclusions

When comparing the above described photo-activation, intra-molecular signal-relay and signal-transduction mechanisms of two eukaryotic (Phototropin: 4.4.1, VVD: 4.4.2) and two prokaryotic (EL222: 4.4.3, YtvA: 4.4.4) LOV photoreceptors to the data presented here for PpSB1-LOV, a number of similarities and differences are apparent.

First of all, in all so far studied LOV systems (including PpSB1-LOV) a highly conserved glutamine residue (Q116 in PpSB1-LOV, see chapter 4.3.1) was suggested to be involved in LOV photo-activation. However, the mode of the structural rearrangement that this residue undergoes upon illumination has yet to be defined experimentally. X-ray crystallographic studies show that the respective glutamine is part of the H-bonding network that coordinates the FMN chromophore within the LOV domain, which stresses the importance of other FMN coordinating residues (such as i.e. N95 in PpSB1-LOV, see chapter 4.3.1) for LOV photo-activation.

A second similarity is that in all described LOV systems (including PpSB1-LOV, chapter 4.3) the structural signal appears to be relayed from the site of photon-capture (FMN chromophore) to N- and C-terminally localized auxiliary α -helical extensions (A' α and J α) (see also chapter 4.1.1).

In contrast, the final structural consequences that within the cell allow for light-dependent signal-transduction appear to be distinct in each so far studied case. While J α -helix dissociation and unfolding is observed for AsLOV2 (4.4.1), the respective segment remains folded in PpSB1-LOV (chapter 4.1.1), YtvA and EL222 (4.4.3). In VVD a corresponding structural element is even completely absent (4.4.2).

From the above described examples YtvA appears to be both structurally and possibly also with respect to photo-activation and signal-relay the closest relative of PpSB1-LOV. Both proteins appear to possess N-terminal cap structures as well as protruding C-terminal $J\alpha$ -helices. The latter ones might even organize into similar coiled-coil like structures in both type of photoreceptor proteins (chapter 4.1.2.3). Hence it is tempting to speculate whether PpSB1-LOV experiences similar quaternary structural changes as YtvA, i.e. a light induced rotation of the two subunits relative to each other. In PpSB1-LOV this would result in a rotational movement of the C-terminal $J\alpha$ -helices or could even trigger complete or partial dissolution of the coiled-coil interaction in one of the two structural states. This notion might be supported by the observation that the isolated $J\alpha$ -peptide of PpSB1-LOV does readily adopt α -helical conformation in solution but apparently does not form a stable enough coiled coil in the absence of the anchoring LOV domain (chapter 4.1.2.3).

Last but not least, with regard to signal-transduction one feasible way for the “effector-less” PpSB1-LOV protein to transduce a signal within the cell, would be via a light-dependent alteration of protein-protein interactions with yet to identify downstream signaling partners. All of the above discussed possibilities (i.e. subunit rotation, $J\alpha$ -helix rotation / $J\alpha$ coiled-coil dissolution, B β -strand folding/unfolding) would be in line with such a signal-transduction mechanism. All of those structural changes could potentially alter PpSB1-LOV surface properties (i.e. electrostatic potential, hydrophobicity) which might result in altered protein-protein interaction affinities within the cell.

4.6 Outlook

The here presented data provides a first glimpse into the nature of PpSB1-LOV photo-activation and intra-molecular signal relay. Yet, without high-resolution X-ray or NMR structures of the protein in both states no unequivocal conclusions about those aspects as well as about potential modes of signal-transduction are possible. The complete backbone resonance assignment for the dark- and light-state of PpSB1-LOV, presented as part of this thesis, represents the first step toward this goal. The here discussed PpSB1-LOV photo-activation and signal relay mechanism represents merely a hypothesis, that however can be tested experimentally i.e. by site-directed mutagenesis of some of the key residues identified in this study, followed by further NMR spectroscopic and other biophysical solution studies.

At the same time the here presented data as well recent X-ray crystallographic evidence [157], highlight the possibility of using short LOV proteins such as PpSB1-LOV as sensory switches for the construction of LOV-based optogenetic tools i.e. in a similar way as the YtvA-LOV domain was used to construct an artificial light-gated kinase (see chapter 4.4.4) [186]. Hereby, they could be used either in *cis*, by fusion of an artificial effector domain module [81, 186] or in *trans* employing a yet to identify regulatory protein [187] or even by using artificially constructed binding proteins as regulatory/interaction modules.

References

1. Grossman, C.L, *A beacon of hope: The importance of light in religion*. USA weekend, 2011.
2. Michelou, V.K., et al., *Light-stimulated bacterial production and amino acid assimilation by cyanobacteria and other microbes in the North Atlantic ocean*. Applied and environmental microbiology, 2007. **73**: p. 5539-46.
3. Wieiss, J, *Some remarks on the photosynthesis of green plants*. Journal of general physiology, 1937. **20**: p. 501-9.
4. Bakalis, N., *Handbook of Greek Philosophy: From Thales to the Stoics Analysis and Fragments*, 2005.
5. Burnet, J., *Early Greek Philosophy*, 1930.
6. Stites, N.P, *Light I: Particle or Wave?*, Visionlearning, 2005.
7. Starr, C., *Biology: Concepts and Applications*, 2005.
8. Pattison, D.I. and M.J. Davies, *Actions of ultraviolet light on cellular structures*. EXS, 2006. **96**: p. 131-57.
9. Liscum, E., et al., *Blue light signaling through the cryptochromes and phototropins. So that's what the blues is all about*. Plant physiology, 2003. **133**: p. 1429-36.
10. Batschauer, A., *Photoreceptors and Light Signalling*, 2003.
11. Spudich, J.L., et al., *Retinylidene proteins: structures and functions from archaea to humans*. Annual review of cell and developmental biology, 2000. **16**: p. 365-92.
12. Hoff, W.D., et al., *Molecular mechanism of photosignaling by archaeal sensory rhodopsins*. Annual review of biophysics and biomolecular structure, 1997. **26**: p. 223-58.
13. Quail, P.H, *The phytochrome family: dissection of functional roles and signalling pathways among family members*. Philosophical transactions of the Royal Society of London. Biological sciences, 1998. **353**: p. 1399-403.
14. Kort, R., et al., *The xanthopsins: a new family of eubacterial blue-light photoreceptors*. The EMBO journal, 1996. **15**(13): p. 3209-18.
15. Wilson, A., et al., *A photoactive carotenoid protein acting as light intensity sensor*. Proceedings of the national academy of sciences of the United States of America, 2008. **105**: p. 12075-80.
16. Rizzini, L., et al., *Perception of UV-B by the Arabidopsis UVR8 protein*. Science, 2011. **332**: p. 103-6.
17. Gomelsky, M. and G. Klug, *BLUF: a novel FAD-binding domain involved in sensory transduction in microorganisms*. Trends in biochemical sciences, 2002. **27**: p. 497-500.
18. Ahmad, M. and A.R. Cashmore, *HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor*. Nature, 1993. **366**: p. 162-6.
19. Huala, E., et al., *Arabidopsis NPH1: a protein kinase with a putative redox-sensing domain*. Science, 1997. **278**: p. 2120-3.
20. Somers, D.E., et al., *ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis*. Cell, 2000. **101**: p. 319-29.
21. Herrou, J. and S. Crosson, *Function, structure and mechanism of bacterial photosensory LOV proteins*. Nature reviews. Microbiology, 2011. **9**: p. 713-23.
22. Losi, A. and W. Gärtner, *Old chromophores, new photoactivation paradigms, trendy applications: flavins in blue light-sensing photoreceptors*. Photochemistry and photobiology, 2011. **87**: p. 491-510.
23. Möglich, A., et al., *Structure and function of plant photoreceptors*. Annual review of plant biology, 2010. **61**: p. 21-47.

24. Losi, A. and W. Gärtner, *The evolution of flavin-binding photoreceptors: an ancient chromophore serving trendy blue-light sensors*. Annual review of plant biology, 2012. **63**: p. 49-72.
25. Chaves, I., et al., *The cryptochromes: blue light photoreceptors in plants and animals*. Annual review of plant biology, 2011. **62**: p. 335-64.
26. Cheng, P., et al., *Functional conservation of light, oxygen, or voltage domains in light sensing*. Proceedings of the national academy of sciences of the United States of America, 2003. **100**: p. 5938-43.
27. Han, Y., et al., *A eukaryotic BLUF domain mediates light-dependent gene expression in the purple bacterium Rhodospirillum rubrum* 2.4.1. Proceedings of the national academy of sciences of the United States of America, 2004. **101**: p. 12306-11.
28. Gauden, M., et al., *Photocycle of the flavin-binding photoreceptor AppA, a bacterial transcriptional antirepressor of photosynthesis genes*. Biochemistry, 2005. **44**: p. 3653-62.
29. Anderson, S., et al., *Structure of a novel photoreceptor, the BLUF domain of AppA from Rhodospirillum rubrum*. Biochemistry, 2005. **44**: p. 7998-05.
30. Masuda, S. and C.E Bauer, *AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in Rhodospirillum rubrum*. Cell, 2002. **110**: p. 613-23.
31. Rajagopal, S., et al., *Purification and initial characterization of a putative blue light-regulated phosphodiesterase from Escherichia coli*. Photochemistry and photobiology, 2004. **80**: p. 542-7.
32. Iseki, M., et al., *A blue-light-activated adenylyl cyclase mediates photoavoidance in Euglena gracilis*. Nature, 2002. **415**: p. 1047-51.
33. Yuan, H., et al., *Mutational and structural studies of the PixD BLUF output signal that affects light-regulated interactions with PixE*. Biochemistry, 2011. **50**: p. 6365-75.
34. Fukushima, Y., et al., *Primary intermediate in the photocycle of a blue-light sensory BLUF FAD-protein, Tll0078, of Thermosynechococcus elongatus BP-1*. Biochemistry, 2005. **44**: p. 5149-58.
35. Okajima, K., et al., *Biochemical and functional characterization of BLUF-type flavin-binding proteins of two species of cyanobacteria*. Journal of biochemistry, 2005. **137**: p. 741-50.
36. Losi, A., *Flavin-based blue-light photosensors: A photobiophysics update*. Photochemistry and photobiology, 2007. **83**: p. 1283-1300.
37. Jung, A., et al., *Crystal structures of the AppA BLUF domain photoreceptor provide insights into blue light-mediated signal transduction*. Journal of molecular biology, 2006. **362**: p. 717-32.
38. Ahmad, M., et al., *Mutations throughout an Arabidopsis blue-light photoreceptor impair blue-light-responsive anthocyanin accumulation and inhibition of hypocotyl elongation*. The Plant journal : for cell and molecular biology, 1995. **8**: p. 653-8.
39. Hitomi, K., et al., *Bacterial cryptochrome and photolyase: characterization of two photolyase-like genes of Synechocystis sp. PCC6803*. Nucleic acids research, 2000. **28**: p. 2353-62.
40. Liu, B., et al., *Searching for a photocycle of the cryptochrome photoreceptors*. Current opinion in plant biology, 2010. **13**: p. 578-86.
41. Bouly, J.P., et al., *Cryptochrome blue light photoreceptors are activated through interconversion of flavin redox states*. Journal of biological chemistry, 2007. **282**: p. 9383-91.
42. Krauss, U., et al., *Distribution and phylogeny of light-oxygen-voltage-blue-light-signaling proteins in the three kingdoms of life*. Journal of bacteriology, 2009. **191**: p. 7234-42.
43. Djouani-Tahri el-B., et al., *A eukaryotic LOV-histidine kinase with circadian clock function in the picoalga Ostreococcus*. The Plant journal : for cell and molecular biology, 2011. **65**: p. 578-88.

44. Purcell, E.B. and S. Crosson, *Photoregulation in prokaryotes*. Current opinion in microbiology, 2008. **11**: p. 168-78.
45. Cao, Z., et al., *A blue light inducible two-component signal transduction system in the plant pathogen Pseudomonas syringae pv. tomato*. Biophysical journal, 2008. **94**: p. 897-905.
46. Swartz, T.E., et al., *Blue-light-activated histidine kinases: two-component sensors in bacteria*. Science, 2007. **317**: p. 1090-3.
47. Purcell, E.B., et al., *A photosensory two-component system regulates bacterial cell attachment*. Proceedings of the national academy of sciences of the United States of America, 2007. **104**: p. 18241-6.
48. Krauss, U., et al., *Initial characterization of a blue-light sensing, phototropin-related protein from Pseudomonas putida: a paradigm for an extended LOV construct*. Physical chemistry chemical physics, 2005. **7**: p. 2804-11.
49. Jentzsch, K., et al., *Mutual exchange of kinetic properties by extended mutagenesis in two short LOV domain proteins from Pseudomonas putida*. Biochemistry, 2009. **48**: p. 10321-33.
50. Avila-Perez, M., et al., *Blue light activates the sigmaB-dependent stress response of Bacillus subtilis via YtvA*. Journal of bacteriology, 2006. **188**: p. 6411-4.
51. Gaidenko, T.A., et al., *The blue-light receptor YtvA acts in the environmental stress signaling pathway of Bacillus subtilis*. Journal of bacteriology, 2006. **188**: p. 6387-95.
52. Losi, A., *The bacterial counterparts of plant phototropins*. Photochemical & photobiological sciences, 2004. **3**: p. 566-74.
53. Losi, A. and W. Gärtner, *Bacterial bilin- and flavin-binding photoreceptors*. Photochemical & photobiological sciences, 2008. **7**: p. 1168-78.
54. Zoltowski, B.D., et al., *Conformational switching in the fungal light sensor Vivid*. Science, 2007. **316**: p. 1054-7.
55. Metz, S., et al., *Role of a short light, oxygen, voltage (LOV) domain protein in blue light- and singlet oxygen-dependent gene regulation in Rhodobacter sphaeroides*. Microbiology, 2012. **158**: p. 368-79.
56. Swartz, T.E., et al., *The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin*. The Journal of biological chemistry, 2001. **276**: p. 36493-500.
57. Matsuoka, D., et al., *Primary Processes During the Light-signal Transduction of Phototropin*. Photochemistry and photobiology, 2007. **83**: p. 470.
58. Richter, G., et al., *Photochemically induced dynamic nuclear polarization in a C450A mutant of the LOV2 domain of the Avena sativa blue-light receptor phototropin*. Journal of the american chemical society, 2005. **127**: p. 17245-52.
59. Losi, A., et al., *First evidence for phototropin-related blue-light receptors in prokaryotes*. Biophysical journal, 2002. **82**: p. 2627-34.
60. Tang, Y., et al., *Interdomain signalling in the blue-light sensing and GTP-binding protein YtvA: a mutagenesis study uncovering the importance of specific protein sites*. Photochemical & photobiological sciences, 2010. **9**: p. 47-56.
61. Kennis, J.T., et al., *Primary reactions of the LOV2 domain of phototropin, a plant blue-light photoreceptor*. Biochemistry, 2003. **42**: p. 3385-92.
62. Salomon, M., et al., *An optomechanical transducer in the blue light receptor phototropin from Avena sativa*. Proceedings of the national academy of sciences of the United States of America, 2001. **98**: p. 12357-61.
63. Zoltowski, B.D., et al., *Mechanism-based tuning of a LOV domain photoreceptor*. Nature chemical biology, 2009. **5**: p. 827-34.
64. Swartz, T.E., et al., *The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin*. Journal of biological chemistry, 2001. **276**: p. 36493-500.
65. Vreede, J., et al., *PAS domains. Common structure and common flexibility*. The Journal of biological chemistry, 2003. **278**: p. 18434-9.

66. Crosson, S. and K. Moffat, *Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction*. Proceedings of the national academy of sciences of the United States of America, 2001. **98**: p. 2995-3000.
67. Fedorov, R., et al., *Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii*. Biophysical journal, 2003. **84**: p. 2474-82.
68. Nambu, J.R., et al., *The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development*. Cell, 1991. **67**: p. 1157-67.
69. Möglich, A., et al., *Structure and signaling mechanism of Per-ARNT-Sim domains*. Structure, 2009. **17**: p. 1282-94.
70. Halavaty, A.S. and K. Moffat, *N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa*. Biochemistry, 2007. **46**: p. 14001-09.
71. Harper, S.M., et al., *Structural basis of a phototropin light switch*. Science, 2003. **301**: p. 1541-4.
72. Vaidya, A.T., et al., *Structure of a Light-Activated LOV Protein Dimer That Regulates Transcription*. Science signaling, 2011. **4**: ra50.
73. Möglich, A. and K. Moffat, *Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA*. Journal of molecular biology, 2007. **373**: p. 112-26.
74. Harper, S.M., et al., *Disruption of the LOV-Jalpha helix interaction activates phototropin kinase activity*. Biochemistry, 2004. **43**: p. 16184-92.
75. Halavaty, A.S. and K. Moffat, *N- and C-terminal flanking regions modulate light-induced signal transduction in the LOV2 domain of the blue light sensor phototropin 1 from Avena sativa*. Biochemistry, 2007. **46**: p. 14001-9.
76. Zayner, J.P., et al., *The amino-terminal helix modulates light-activated conformational changes in AsLOV2*. Journal of molecular biology, 2012. **419**: p. 61-74.
77. Strickland, D., et al., *TULIPs: tunable, light-controlled interacting protein tags for cell biology*. Nature methods, 2012. **9**: p. 379-84.
78. Lee, J., et al., *Surface sites for engineering allosteric control in proteins*. Science, 2008. **322**: p. 438-42.
79. Strickland, D., et al., *Light-activated DNA binding in a designed allosteric protein*. Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**: p. 10709-14.
80. Wu, Y.I., et al., *A genetically encoded photoactivatable Rac controls the motility of living cells*. Nature, 2009. **461**: p. 104-8.
81. Ohlendorf, R., et al., *From Dusk till Dawn: One-Plasmid Systems for Light-Regulated Gene Expression*. Journal of molecular biology, 2012. **416**: p. 534-42.
82. Lungu, O.I., et al., *Designing Photoswitchable Peptides Using the AsLOV2 Domain*. Chemistry & biology, 2012. **19**: p. 507-517.
83. Schierling, B. and A. Pingoud, *Controlling the DNA Cleavage Activity of Light-Inducible Chimeric Endonucleases by Bidirectional Photoactivation*. Bioconjugate chemistry, 2012. **23**: p. 1105-09.
84. Christie, J.M., et al., *LOV to BLUF: Flavoprotein Contributions to the Optogenetic Toolkit*. Molecular Plant, 2012. **5**: p. 533-44.
85. Purcell, E.M., et al., *Resonance Absorption by Nuclear Magnetic Moments in a Solid*. Physical review, 1946. **69**: p. 37-8.
86. Bloch, F., et al., *Nuclear Induction*. Physical Review, 1946. **69**: p. 127.
87. Williamson, M.P., et al., *Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ¹H nuclear magnetic resonance and distance geometry*. Journal of molecular biology, 1985. **182**: p. 295-315.

88. Ernst, R.R. and W.A. Anderson, *Sensitivity Enhancement in Magnetic Resonance. II. Investigation of Intermediate Passage Conditions*. Review of scientific instruments, 1965. **36**: 1696.
89. Wagner, G. and K. Wüthrich, *Dynamic model of globular protein conformations based on NMR studies in solution*. Nature, 1978. **275**: p. 247-8.
90. Markwick, P.R.L., et al., *Structural Biology by NMR: Structure, Dynamics, and Interactions*. Plos computational biology, 2008. **4**.
91. Mittermaier, A. and L.E. Kay, *New tools provide new insights in NMR studies of protein dynamics*. Science, 2006. **312**: p. 224-8.
92. Pervushin K., et al, *Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution*. Proceedings of the national academy of sciences of the United States of America, 1997. **94**: p. 12366-71.
93. Becker, E.D. and T.C. Farrar, *Fourier transform spectroscopy*. Science, 1972. **178**: p. 361-8.
94. Levitt, M.H., *Spin Dynamics: Basics of Nuclear Magnetic Resonance*, 2009.
95. Cheng, H., et al., *Protein expression, selective isotopic labeling, and analysis of hyperfine-shifted NMR signals of Anabaena 7120 vegetative [2Fe-2S]ferredoxin*. Archives of biochemistry and biophysics, 1995. **316**: p. 619-34.
96. Bax, A. and M. Ikura, *An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins*. Journal of biomolecular NMR, 1991. **1**: p. 99-104.
97. Fischer, M.W., et al., *Domain orientation and dynamics in multidomain proteins from residual dipolar couplings*. Biochemistry, 1999. **38**: p. 9013-22.
98. Korzhnev, D.M., et al., *Backbone dynamics of the channel-forming antibiotic zervamicin IIB studied by 15N NMR relaxation*. FEBS letters, 2001. **495**: p. 52-5.
99. Hamuro, Y., et al., *Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry*. Journal of biomolecular techniques, 2003. **14**: p. 171-82.
100. Crosson, S. and K. Moffat, *Photoexcited structure of a plant photoreceptor domain reveals a light-driven molecular switch*. The Plant cell, 2002. **14**: p. 1067-75.
101. Fedorov, R., et al., *Crystal structures and molecular mechanism of a light-induced signaling switch: The Phot-LOV1 domain from Chlamydomonas reinhardtii*. Biophysical journal, 2003. **84**: p. 2474-82.
102. Studier, F.W. and B.A. Moffatt, *Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes*. Journal of molecular biology, 1986. **189**: p. 113-30.
103. Woodcock, D.M., et al., *Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants*. Nucleic acids research, 1989. **17**: p. 3469-78.
104. Mathes, T., et al., *In vivo generation of flavoproteins with modified cofactors*. Journal of molecular biology, 2009. **385**: p. 1511-8.
105. Rieping, W., et al., *ARIA2: automated NOE assignment and data integration in NMR structure calculation*. Bioinformatics, 2007. **23**: p. 381-2.
106. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline*. Proteins, 2005. **59**: p. 687-96.
107. Thompson, J.D., et al., *CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice*. Nucleic acids research, 1994. **22**: p. 4673-80.
108. Doreleijers, J.F., et al., *NRG-CING: integrated validation reports of remediated experimental biomolecular NMR data and coordinates in wwPDB*. Nucleic acids research, 2012. **40**: p. D519-24.

109. Brunger, A.T., et al., *Crystallography & NMR system: A new software suite for macromolecular structure determination*. Acta crystallographica section D: Biological crystallography, 1998. **54**: p. 905-21.
110. Gasteiger, E., et al., *ExpASY: The proteomics server for in-depth protein knowledge and analysis*. Nucleic acids research, 2003. **31**: p. 3784-8.
111. Davis, I.W., et al., *MolProbity: all-atom contacts and structure validation for proteins and nucleic acids*. Nucleic acids research, 2007. **35**: p. 375-83.
112. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. Journal of biomolecular NMR, 1995. **6**: p. 277-93.
113. DeLano, W.L., *The PyMOL Molecular Graphics System*. on World Wide Web <http://www.pymol.org>, 2002.
114. Shen, Y., et al., *TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts*. Journal of biomolecular NMR, 2009. **44**: p. 213-23.
115. Green, M.R. and J. Sambrook, *Molecular cloning: A laboratory manual*, 1989.
116. Studier, F.W., *Protein production by auto-induction in high density shaking cultures*. Protein expression and purification, 2005. **41**: p. 207-34.
117. Sivashanmugam, A., et al., *Practical protocols for production of very high yields of recombinant proteins using Escherichia coli*. Protein science, 2009. **18**: p. 936-48.
118. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*. Nature, 1970. **227**: p. 680-5.
119. Dyballa, N. and S. Metzger, *Fast and Sensitive Colloidal Coomassie G-250 Staining for Proteins in Polyacrylamide Gel*. Journal of visualized experiments, 2009. **30**: p. 1431.
120. Morrison, D.A., *Transformation in Escherichia coli: cryogenic preservation of competent cells*. Journal of bacteriology, 1977. **132**: p. 349-51.
121. Roberts, R.J., *Restriction endonucleases*. Critical review in biochemistry, 1976. **2**: p.123-64.
122. Lehman, I.R., *DNA ligase: structure, mechanism, and function*. Science, 1974. **186**: p. 790-7.
123. Mullis, K., et al., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction*. Cold Spring Harbor symposia on quantitative biology, 1986. **51**: p. 263-73.
124. Sanger, F. and A.R. Coulson, *A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase*. Journal of molecular biology, 1975. **94**: p. 441-8.
125. Di Cera, et al, *Determinants of thrombin specificity*. Annals of the New York Academy of Sciences, 2001. **936**: p. 133-46.
126. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. Analytical biochemistry, 1976. **72**: p. 248-54.
127. Lathe, G.H. and C.R Ruthven, *The separation of substances on the basis of their molecular weights, using columns of starch and water*. Biochemical journal, 1955. **60**.
128. Nakasone, Y., et al., *Kinetic measurement of transient dimerization and dissociation reactions of Arabidopsis phototropin 1 LOV2 domain*. Biophysical journal, 2006. **91**: p. 645-53.
129. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. Analytical chemistry, 1988. **60**: p. 2299-301.
130. Perczel, A., et al., *Convex constraint analysis: a natural deconvolution of circular dichroism curves of proteins*. Protein engineering, 1991. **4**: p. 669-79.
131. Altschul, S.F. and E.V. Koonin, *Iterated profile searches with PSI-BLAST--a tool for discovery in protein databases*. Trends in biochemical sciences, 1998. **23**: p. 444-7.
132. Schultz, J., et al., *SMART, a simple modular architecture research tool: identification of signaling domains*. Proceedings of the national academy of sciences of the United States of America, 1998. **95**: p. 5857-64.
133. Notredame, C., et al., *T-Coffee: A novel method for fast and accurate multiple sequence alignment*. Journal of molecular biology, 2000. **302**: p. 205-17.

134. Gruber, M., et al., *Comparative analysis of coiled-coil prediction methods*. Journal of structural biology, 2006. **155**: p. 140-5.
135. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein*. Journal of molecular biology, 1982. **157**: p. 105-32.
136. Bax, A. and D.G. Davis, *Practical Aspects of Two-Dimensional Transverse Noe Spectroscopy*. Journal of magnetic resonance, 1985. **63**: p. 207-13.
137. Jeener, J., et al., *Investigation of Exchange Processes by 2-Dimensional Nmr-Spectroscopy*. Journal of chemical physics, 1979. **71**: p. 4546-53.
138. Wüthrich, K., *NMR of Proteins and Nucleic Acids*, 1986.
139. Bodenhausen, G. and D.J. Ruben, *Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy*. Chemical Physics Letters 1980. **69**: p. 185-9.
140. Kay, L.E., et al., *Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins*. Journal of magnetic resonance, 1990. **89**: p. 496-514.
141. Bax, A. and M. Ikura, *An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins*. Journal of biomolecular NMR, 1991. **1**: p. 99-104.
142. Robert, T.C., et al., *A constant-time three-dimensional triple-resonance pulse scheme to correlate intraresidue 1HN, 15N, and 13C' chemical shifts in 15N--13C-labelled proteins*. Journal of Magnetic Resonance, 1992. **97**: p. 213-7.
143. Zuiderweg, E.R. and S.W. Fesik, *Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a*. Biochemistry, 1989. **28**: p. 2387-91.
144. Farmer, B.T., et al., *Localizing the NADP+ binding site on the MurB enzyme by NMR*. Nature structural biology, 1996. **3**: p. 995-7.
145. Koradi, R., et al., *MOLMOL: a program for display and analysis of macromolecular structures*. Journal of molecular graphics, 1996. **14**: p. 51-5.
146. Vriend, G., *WHAT IF: a molecular modeling and drug design program*. Journal of molecular graphics , 1990 Mar. **8**: p. 52-6.
147. Hendrischk, A.K., et al., *Characterization of an unusual LOV domain protein in the alpha-proteobacterium Rhodobacter sphaeroides*. Photochemistry and photobiology, 2009. **85**: p. 1254-9.
148. Bhagwat, M. and L. Aravind, *PSI-BLAST tutorial*. Methods in molecular biology, 2007. **395**: p. 177-86.
149. Moutevelis, E. and D.N. Woolfson, *A periodic table of coiled-coil protein structures*. Journal of molecular biology, 2009. **385**: p. 726-32.
150. Brahm, S. and J. Brahm, *Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism*. Journal of molecular biology, 1980. **138**: p. 149-78.
151. Sonnichsen, F.D., et al., *Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide*. Biochemistry, 1992. **31**: p. 8790-8.
152. Najbar, L.V., et al., *Conformational analysis of LYS(11-36), a peptide derived from the beta-sheet region of T4 lysozyme, in TFE and SDS*. Biochemistry, 1997. **36**: p. 11525-33.
153. Marsh, J.A., et al., *Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: implications for fibrillation*. Protein science, 2006. **12**: p. 2795-804.
154. Vogl, C., et al., *Characterization of riboflavin (vitamin B2) transport proteins from Bacillus subtilis and Corynebacterium glutamicum*. Journal of bacteriology, 2007. **189**: p. 7367-75.
155. Pielers, U., et al., *Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a powerful tool for the mass and sequence analysis of natural and modified oligonucleotides*. Nucleic acids research, 1993. **21**: p. 3191-6.

-
-
156. Lee, Y.H., et al., *Cores and pH-dependent dynamics of ferredoxin-NADP⁺ reductase revealed by hydrogen/deuterium exchange*. The Journal of biological chemistry, 2007. **282**: p. 5959-67.
 157. Circolone, F., et al., *Structural basis for the slow dark recovery of a full-length LOV protein from Pseudomonas putida*. Journal of molecular biology, 2012. **417**: p. 362-74.
 158. Buttani, V., et al., *Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions*. Photochemical & photobiological sciences : Official journal of the european photochemistry association and the european society for photobiology, 2007. **6**: p. 41-9.
 159. Key, J. and K. Moffat, *Crystal structures of deoxy and CO-bound bFixLH reveal details of ligand recognition and signaling*. Biochemistry, 2005. **44**: p. 4627-35.
 160. Key, J., et al., *Structure of the redox sensor domain of Azotobacter vinelandii NifL at atomic resolution: Signaling, dimerization, and mechanism*. Biochemistry, 2007. **46**: p. 3614-23.
 161. Sevana, M., et al., *A ligand-induced switch in the periplasmic domain of sensor histidine kinase CitA*. Journal of molecular biology, 2008. **377**: p. 512-23.
 162. Park, H.J., et al., *Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of Escherichia coli dos heme domain (EcDosH)*. Biochemistry, 2004. **43**: p. 2738-46.
 163. Miyatake, H., et al., *Sensory mechanism of oxygen sensor FixL from Rhizobium meliloti: Crystallographic, mutagenesis and resonance Raman spectroscopic studies*. Journal of molecular biology, 2000. **301**: p. 415-31.
 164. Ma, X.L., et al., *PAS-mediated dimerization of soluble guanylyl cyclase revealed by signal transduction histidine kinase domain crystal structure*. Journal of biological chemistry, 2008. **283**: p. 1167-78.
 165. Apostolovic, B., et al., *Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials*. Chemical society reviews, 2010. **39**: p. 3541-75.
 166. Hill, A.F., *Prion protein protocols*, 2008.
 167. Dutta, K., et al., *pH-induced folding of an apoptotic coiled coil*. Protein science 2001. **10**: p. 2531-40.
 168. Muhle-Goll, C., et al., *The dimerization stability of the HLH-LZ transcription protein family is modulated by the leucine zippers: a CD and NMR study of TFEB and c-Myc*. Biochemistry, 1994. **33**: p. 11296-306.
 169. Comeau, S.R., et al., *ClusPro: a fully automated algorithm for protein-protein docking*. Nucleic acids research, 2004. **32**: p. 96-9.
 170. Kabsch, W. and C. Sander, *Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features*. Biopolymers, 1983. **22**: p. 2577-637.
 171. Avila-Perez, M., et al., *In vivo mutational analysis of YtvA from Bacillus subtilis: mechanism of light activation of the general stress response*. The Journal of biological chemistry, 2009. **284**: p. 24958-64.
 172. Nash, A.I., et al., *A conserved glutamine plays a central role in LOV domain signal transmission and its duration*. Biochemistry, 2008. **47**: p. 13842-9.
 173. Zoltowski, B.D. and B.R. Crane, *Light activation of the LOV protein vivid generates a rapidly exchanging dimer*. Biochemistry, 2008. **47**: p. 7012-9.
 174. Rinaldi, J., et al., *The β -scaffold of the LOV domain of the Brucella light-activated histidine kinase is a key element for signal transduction*. Journal of molecular biology, 2012. **420**: p. 112-27.
 175. Krieger, E., et al, *Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field*. Proteins, 2002. **47**: p. 393-402.

-
-
176. Yao, X., et al., *Estimation of the available free energy in a LOV2-J alpha photoswitch*. Nature chemical biology, 2008. **4**: p. 491-7.
 177. Nash, A.I., et al., *Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein*. Proceedings of the national academy of sciences of the United States of America, 2011. **108**: p. 9449-54.
 178. Wishart, D.S., et al., *The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy*. Biochemistry, 1992. **31**: p. 1647-51.
 179. Berjanskii, M.V. and D.S. Wishart, *The RCI server: rapid and accurate calculation of protein flexibility using chemical shifts*. Nucleic acids research, 2007. **35**: p. 531-7.
 180. Jurk, M., et al., *Blue flickers of hope: secondary structure, dynamics, and putative dimerization interface of the blue-light receptor YtvA from Bacillus subtilis*. Biochemistry, 2011. **50**: p. 8163-71.
 181. Sakamoto, K. and W.R. Briggs, *Cellular and subcellular localization of phototropin 1*. The Plant cell, 2002. **14**: p. 1723-35.
 182. Briggs, W.R. and J.M. Christie, *Phototropins 1 and 2: versatile plant blue-light receptors*. Trends in plant science, 2002. **7**: p. 204-10.
 183. Chan, R.H., et al., *Photocycle of the LOV-STAS Protein from the Pathogen Listeria monocytogenes*. Photochemistry and photobiology, 2012.
 184. Losi, A., et al., *Mutational effects on protein structural changes and interdomain interactions in the blue-light sensing LOV protein YtvA*. Photochemistry and photobiology, 2005. **81**: p. 1145-52.
 185. Jurk, M., et al., *The switch that does not flip: the bluelight receptor YtvA from Bacillus subtilis adopts an elongated dimer conformation independent of the activation state as revealed by a combined AUC and SAXS study*. Journal of molecular biology, 2010. **403**: p. 78-87.
 186. Möglich, A., et al., *Design and signaling mechanism of light-regulated histidine kinases*. Journal of molecular biology, 2009. **385**: p. 1433-44.
 187. Yazawa, M., et al., *Induction of protein-protein interactions in live cells using light*. Nature biotechnology, 2009. **27**: p. 941-5.

Appendix A

Table A.1: Acquisition and processing parameters of NMR spectroscopic experiments

Experiments	sw	sw ₁	sw ₂	t ₁	t ₂	t ₃	scans	d1	mix	¹⁵ N	¹³ C	Ω ₁	Ω ₂	Ω ₃
	(ppm)	(ppm)	(ppm)	(pt)	(pt)	(pt)		(s)	(ms)	(ppm)	(ppm)	(pt)	(pt)	(pt)
2D (¹ H- ¹⁵ N)-HSQC	13.35	40.00	-	1022	128	-	8	1	-	119.35	-	4096	2048	-
2D (¹ H- ¹ H)TOCSY	10.06	10.06	-	1812	256	-	32	1	60	-	-	4096	2048	-
2D (¹ H- ¹ H)NOESY	10.06	10.06	-	1812	256	-	64	1	350	-	-	4096	2048	-
3D HNCQ	13.35	16.65	37.28	1024	70	60	16	1	-	117.19	176.5	4096	256	256
3D HNCACO	13.35	16.65	37.28	1024	50	44	32	1	-	117.19	176.5	4096	256	256
3D HNCA	13.35	34.95	35.09	1022	100	48	32	0.8	-	117.19	55.88	4096	512	256
3D HNCOCA	13.35	34.95	35.09	1024	100	48	32	0.8	-	117.19	55.88	2048	512	256
3D HCCH-TOCSY	13.35	7.77	80	1022	60	96	8	1.8	-	-	43	2048	512	256
3D (¹ H- ¹ H- ¹⁵ N)-NOESY-HSQC	12.29	12.22	35.09	940	200	60	8	1.2	-	117.19	-	4048	512	256
3D (¹ H- ¹³ C- ¹ H)-HSQC-NOESY	13.35	6.66	80	1082	110	76	16	1	-	-	40	4048	512	256
2D (¹ H) ¹⁵ N-NOE	16.67	41.11	-	850	128	-	32	3	-	118.12	-	8192	1024	-
2D T1-Relaxation	16.67	41.11	-	852	96	-	8	5	-	118.12	-	8192	1024	-
2D T2-Relaxation	16.67	41.11	-	852	96	-	8	6	-	118.12	-	8192	1024	-

sw: acquired spectral width in the respective dimension t: acquired number of points in the respective dimension
d1: relaxation delay time mix: NOESY/TOCSY mixing/spinlock time
¹⁵N or ¹³C carrier frequencies in ppm Ω: number of real points in the respective dimension after Fourier transform

Table A.2: Short LOV proteins with Y(Q/R)DCRFLQG motif identified in the prokaryotic kingdom

Tree-ID	Organims	UniProt/Genbank	Subclass	Order
PpSB1-LOV	<i>Pseudomonas putida</i> KT2440	Q88E39	Gamma	<i>Pseudomonadales</i>
PpSB2-LOV	<i>Pseudomonas putida</i> KT2440	Q88JB0	Gamma	<i>Pseudomonadales</i>
Pput_F1-1	<i>Pseudomonas putida</i> F1	A5W8Z9	Gamma	<i>Pseudomonadales</i>
Pput_F1-2	<i>Pseudomonas putida</i> F1	A5W4T2	Gamma	<i>Pseudomonadales</i>
Pput_S16-1	<i>Pseudomonas putida</i> S16	F8G5Y7	Gamma	<i>Pseudomonadales</i>
Pput_S16-2	<i>Pseudomonas putida</i> S16	F8G2U7	Gamma	<i>Pseudomonadales</i>
Pput_W619_1	<i>Pseudomonas putida</i> W619	B1J385	Gamma	<i>Pseudomonadales</i>
PputW619_2	<i>Pseudomonas putida</i> W619	B1JAC4	Gamma	<i>Pseudomonadales</i>
Ps_TJI-51-1	<i>Pseudomonas</i> sp. TJI-51	ZP_08139218.1	Gamma	<i>Pseudomonadales</i>
Ps_TJI-51-2	<i>Pseudomonas</i> sp. TJI-51	ZP_08142158.1	Gamma	<i>Pseudomonadales</i>
Pp_BIRD_1	<i>Pseudomonas putida</i> BIRD-1	E4RI35	Gamma	<i>Pseudomonadales</i>
Pp_BIRD_2	<i>Pseudomonas putida</i> BIRD-1	E4R4W4	Gamma	<i>Pseudomonadales</i>
Pput_GB1	<i>Pseudomonas putida</i> GB1	B0KGV4	Gamma	<i>Pseudomonadales</i>
Pflu_Pf-5	<i>Pseudomonas fluorescens</i> Pf-5	Q4KI48	Gamma	<i>Pseudomonadales</i>
P_ful_12-X	<i>Pseudomonas fulva</i> 12-X	F6AEJ4	Gamma	<i>Pseudomonadales</i>
Pmend_YMP	<i>Pseudomonas mendocina</i> ymp	A4XXH1	Gamma	<i>Pseudomonadales</i>
Pmend_NK-01	<i>Pseudomonas mendocina</i> NK-01	F4DSG2	Gamma	<i>Pseudomonadales</i>
P_braNFM421	<i>Pseudomonas brassicacearum</i> subsp. <i>brassicacearum</i> NFM421	F2KL83	Gamma	<i>Pseudomonadales</i>
Pstu_A1501	<i>Pseudomonas stutzeri</i> A1501	A4VKZ3	Gamma	<i>Pseudomonadales</i>
Pstu_LMG11199	<i>Pseudomonas stutzeri</i> ATCC 17588 = LMG 11199	F8H9C5	Gamma	<i>Pseudomonadales</i>
Pflu_WH6	<i>Pseudomonas fluorescens</i> WH6	ZP_07777456.1	Gamma	<i>Pseudomonadales</i>
Pflu_SBW25	<i>Pseudomonas fluorescens</i> SBW25	C3K1W0	Gamma	<i>Pseudomonadales</i>

Tree-ID	Organims	UniProt/Genbank	Subclass	Order
P_flu-Pf0-1	<i>Pseudomonas fluorescens</i> Pf0-1	Q3KHW7	Gamma	<i>Pseudomonadales</i>
PsF113	<i>Pseudomonas fluorescens</i> F113		Gamma	<i>Pseudomonadales</i>
Csalex	<i>Chromohalobacter salexigens</i> DSM 3043	Q1QU87	Gamma	<i>Oceanospirillales</i>
HTD01	<i>Halomonas</i> sp. TD01	ZP_08635475.1	Gamma	<i>Oceanospirillales</i>
HboLC1	<i>Halomonas boliviensis</i> LC1	ZP_09188734.1	Gamma	<i>Oceanospirillales</i>
HGFAJ-1	<i>Halomonas</i> sp. GFAJ-1	ZP_09288308.1	Gamma	<i>Oceanospirillales</i>
HHAL1	<i>Halomonas</i> sp. HAL1	ZP_08961366.1	Gamma	<i>Oceanospirillales</i>
TcycDSM	<i>Thioalkalimicrobium cyclicum</i> (strain DSM 14477 / JCM 11371 / ALM1)	F6D9J9	Gamma	<i>Thiotrichales</i>
Hnea_ATCC	<i>Halothiobacillus neapolitanus</i>	D0L135	Gamma	<i>Chromatiales</i>
T_K90mix	<i>Thioalkalivibrio</i> sp. (strain K90mix)	D3S9U5	Gamma	<i>Chromatiales</i>
T_EbGR7	<i>Thioalkalivibrio</i> sp. (strain HL-EbGR7)	B8GRG0	Gamma	<i>Chromatiales</i>

Table A.3: Chemical shifts values for PpSB1-J α

Residue	H	Ha	Hb	Hd	He	Hg
119ACE		2.06				
120S	8.25	4.41	3.85,3.92			
121R	8.55	4.3	1.80,1.89	3.17,3.24	7.32	1.65,1.80
122Q	8.36	4.14	2.06,2.06		6.76,7.66	2.36,2.36
123V	7.86	3.88	2.09			0.94,1.00
124E	8.2	4.13	2.07,2.07			2.31,2.31
125L	8.12	4.19	1.62,1.73	0.87,0.89		1.65
126E	8.24	4.01	2.06,2.10			2.27,2.42
127R	8.04	4.13	1.90,1.94	3.18,3.20	7.38	1.62,1.74
128E	8.16	4.08	2.08,2.12			2.26,2.42
129L	8.24	4.13	1.55,1.76	0.87,0.88		1.73
130A	7.99	4.08	1.48			
131E	8.01	4.09	2.09,2.09			2.26,2.40
132L	7.88	4.16	1.64,1.81	0.88,0.92		1.75
133R	7.96	4.16	1.82,1.89	3.16,3.17	7.22	1.64,1.76
134A	7.67	4.31	1.42			
135R	7.79	4.57	1.78,1.86	3.20,3.22	7.51	1.72,1.78
136P		4.43	1.85,2.28	3.61,3.81		2.00,2.00
137K	8.41	4.59	1.72,1.85	1.71,1.71	3.02,3.02	1.48,1.52
138P		4.37	1.91,2.28	3.66,3.81		2.01,2.02
139D	8.34	4.54	2.62,2.68			
140E	8.24	4.28	1.94,2.07			2.23,2.24
141R	8.25	4.32	1.76,1.88	3.20,3.22	7.32	1.63,1.74
142A	7.87	4.12	1.32			

Table A.4: Chemical shifts values for PpSB2-J α

Residue	H	Ha	Hb	Hd	He	Hg
119ACE		2.09				
120T	7.89	4.27	4.25			1.23
121A	8.42	4.28	1.43			
122Q	8.22	4.19	2.01,2.01		6.75,7.60	2.32,2.32
123V	7.77	3.95	2.01			0.84,0.91
124F	8.05	4.54	3.04,3.21	7.26	7.32	
125A	8.14	4.16	1.46			
126E	8.39	3.96	2.06,2.11			2.29,2.41
127E	8.38	4.08	2.07,2.11			2.33,2.37
128R	7.92	4.2	1.90,1.98	3.12,3.31	7.27	1.63,1.68
129V	7.98	3.65	2.18			0.94,1.04
130R	7.93	4.11	1.91,1.97	3.23,3.23	7.52	1.68,1.83
131E	8.07	4.1	2.14,2.23			2.30,2.48

Residue	H	Ha	Hb	Hd	He	Hg
132L	8.18	4.21	1.66,1.88	0.89,0.92		1.73
133E	8.66	3.91	2.02,2.19			2.21,2.54
134A	7.88	4.21	1.56			
135E	8.08	4.1	2.18,2.26			2.28,2.46
136V	8.51	3.62	2.14			0.93,1.05
137A	8.07	4.05	1.54			
138E	7.85	4.31	2.11,2.17			2.26,2.44
139L	8.25	4.08	1.61,1.91	0.87,0.93		1.77
140R	8.51	4.01	1.89,1.92	3.13,3.13	7.16	1.62,1.88
141R	8	4.08	2.02,2.02	3.19,3.29	7.8	1.56,1.77
142Q	8.18	4.08	2.16,2.23		6.67,7.83	2.43,2.62
143Q	8.11	4.2	2.14,2.18		6.69,7.39	2.42,2.54
144G	8.05	3.94,4.01				
145Q	7.88	4.33	2.02,2.17		6.80,7.50	2.40,2.41
146A	8.04	4.31	1.39			
147K	8.12	4.28	1.73,1.83	1.68,1.73	3.00,3.00	1.40,1.46
148H	7.94	4.46	3.07,3.23	7.17	8.38	

Table A.5: Chemical shifts values for PpSB1-LOV dark-state

Residue	H	N	C	Ca
8Q	8.14	117.69	178.22	58.69
9S	7.99	115.79	176.43	62.99
10M	8.13	122.3	176.91	60.61
11V	7.6	117.77	179.43	67.02
12D	8.5	119.97	177.78	57.06
13A	7.68	122.43	176.31	52.55
14S	6.89	110.56	174.13	59.76
15N	9.07	123.14	176.06	55.13
16D	7.91	119.81	178.57	53.75
17G	9.27	110.01	172.42	46.03
18I	8.07	120.69	175.22	61.61
19V	9.3	116.53	175.05	56.88
20V	8.53	117.91	174.27	60.92
21A	9.53	129.84	174.71	50.08
22E	8.98	120.23	174.74	54.05
23K	9.03	125.78	175.93	57.99
24E	8.81	129.82	175.37	55.69
25G	8.78	116.44	174.92	47.05
26D	8.9	127.32	174.5	55.15
27D	7.83	119.39	175.17	53.05
28T	9.94	120.69	174.94	61.3
29I	8.57	119.45	176.22	60.53

Residue	H	N	C	Ca
30L	8.05	122.32	177.69	55.54
31I	9.01	113.74	175.3	60.49
32Y	7.54	121.93	173.87	59.67
33V	7.16	119.08	172.76	59.03
34N	8.01	120.14	174.96	50.73
35A	8.49	121.42	180.89	54.84
36A	7.99	118.57	179.84	55.15
37F	7.93	118.37	178.44	62.88
38E	7.53	120.51	179.65	59.08
39Y	7.97	119.78	178.12	60.08
40L	8.72	119.48	177.2	57.36
41T	7.44	101.7	174.78	63.28
42G	7.8	109.46	173.44	45.36
43Y	7.74	119.27	174.51	58.01
44S	8.72	114.4	176.8	57
45R	9.83	124.22	176.45	59.31
46D	8.09	114.83	177.7	56.6
47E	7.79	116.41	177.81	57.73
48I	6.93	107.81	174.8	59.59
49L	7.19	119.2	177.46	56.54
50Y	8.84	113.91	174.18	58.82
51Q	7.54	115.01	175.42	53.8
52D	8.65	122.86	177.3	54.1
53C	8.79	124.18	174.55	61.45
54R	9.65	120.28	178.16	58.61
55F	7.12	120.48	178.17	59.47
56L	7.33	124.15	177.83	56.75
57Q	7.76	116.21	175.76	57.28
58G	7.63	104.98	174.47	46.73
59D	8.24	124.75	175.97	53.79
60D	8.52	125.58	176.79	52.5
61R	7.65	119.83	176.29	56.06
62D	8.94	123.66	175.74	53.11
63Q	7.33	118.43	178.19	56.43
64L			179.25	58.01
65G	9.06	103.99	175.18	47.14
66R	7.49	118.47	177.58	60.8
67A	7.42	119.25	180.81	55
68R	7.15	116.94	179.51	60.15
69I	8.13	122.14	177.3	65.09
70R	7.99	118.03	179.55	60.47
71K	7.65	118.96	177.54	59.19
72A	7.8	121.21	180.51	55.76
73M	8.48	115.43	179.69	59.53

Residue	H	N	C	Ca
74A	7.84	121.94	179.79	54.78
75E	7.76	115.21	176.67	56.06
76G	8.03	109.87	174	46.79
77R	8.37	119.69	175.33	52.98
78P				
79C	8.44	113.39	171.97	56.52
80R	8.3	124.44	174.81	55.4
81E	8.96	125.47	173.92	53.92
82V	8.42	120.92	176.2	61.68
83L	9.25	128.8	175.32	53.01
84R	9.47	123.16	176.32	55.94
85N	8.48	129.66	170.84	49.8
86Y	8.68	115.65	176.45	57.34
87R	9.17	120.95	178.18	55.75
88K	8.29	124.24	177.35	59.69
89D	7.76	113.45	177.09	53.67
90G	8.32	109.45	174.62	44.92
91S	8.42	117.63	173.02	59.04
92A	8.25	124.36	177.63	50.75
93F	8.8	118.2	172.46	55.66
94W	8.51	122.47	173.03	56.38
94W				
95N	8.96	124.57	175.34	50.53
96E	9.7	127.06	172.4	56.63
97L	8.79	134.22	175.41	53.02
98S	9.08	120.82	174.38	55.98
99I	8.78	124.92	175.67	59.63
100T	9.28	123.39	172.15	
101P				
102V	8.6	122.23	174.09	61.12
103K	8.49	127.17	176.11	55.45
104S	8.61	119.85	175.34	56.92
105D	8.79	125.23	176.9	55.33
106F	8.07	117.78	175.67	58.84
107D	7.64	117.61	175.65	53.16
108Q	8.03	114.51	175.54	56.57
109R	7.98	120.18	175.06	55.71
110T	8.23	117.65	172.99	62.45
111Y	8.77	124.86	174.1	56.3
112F	9.05	116.53	174.38	57.21
113I	8.97	124.18	176.08	59.64
114G	9.26	116.53	169.91	43.87
115I	8.89	124.21	176.08	58.55
116Q	8.93	127.21	173.59	56.58

Residue	H	N	C	Ca
117K	8.16	124.86	175.41	53.97
118D	8.84	125.17	178.14	53.96
119V	8.97	123.07	176.18	60.58
120S	9.18	118.67	175.49	64.06
121R	8.72	119.77	177.7	58.64
122Q	7.83	118.76	177.24	60.02
123V	8	118.86	178.93	66.47
124E	8.11	119.38	179.75	59.22
125L	8.41	121.17	178.75	58.13
126E	8.24	117.74	179.95	60.06
127R	7.97	120.63	178.57	59.4
128E	8.2	121.26	179.33	59.69
129L	8.44	119.48	178.26	58.11
130A	7.85	119.98	180.88	55.07
131E	7.8	117.43	179.05	58.97
132L	7.97	119.44	179.55	57.22
133R	8.2	117.76	176.61	57.97
134A	7.43	121.02	177.3	52.42
135R	7.67	120.42	174.1	54.41
137K	8.38	122.55	174.93	54.35
139D	8.31	119.42		54.24
140E	8.2	121.35		56.64
141R	8.23	122.01		55.93
142A	7.85	130.92		53.91

Table A.6: Chemical shifts values for PpSB1-LOV light-state

Residue	H	N	C	Ca	Cb	Cg	Cd	Ha	Hb	Hd	He	Hg
5Q	8.16	117.61	178.64	59.47	27.93	34.46		3.87	2.03,2.23		6.77,7.45	2.31,2.33
6L	7.83	122.17	179.25	57.69	41.67	27.38	22.89,24.72	4.04	1.30,1.42	0.43,0.51		1.11
7L	7.83	119.6	178.21	58.24	41.06	27.76	22.83,24.42	3.68	1.31,1.63	0.39,0.41		1.22
8Q	8.25	118.62	177.16	59.68	27.99	33.56		3.67	2.01,2.08		6.09,7.33	2.12,2.28
9S	7.8	114.93	176.59	62.25	63.01			4.35	3.99,4.10			7.35
10M	7.88	119.99	177.75	56.94	27.73	31.66		4.4	1.73,1.82		1.59	2.46,2.76
11V	8.15	120.76	178.58	67.31	31.71	21.53,22.88		3.47	2.07			0.81,1.12
12D	8.81	118.28	179.06	57.27	41.24			4.67	2.78,2.99			
13A	7.34	119.57	176.64	52.36	21.52			4.56	1.59			
14S	7.36	111.69	173.76	60.02				3.89				
15N	8.87	124.22	173.97	54.67	40.27			4.63	2.56,2.67	6.74,7.81		
16D	8.1	119.23	174.97	53.58	43.27			4.94	2.32,2.52			
17G	8.63	108.15	172.04	47.07				3.32,4.61				
18I	8.64	130.62	174.16	61.17	41.48	16.33,28.02		5.2	1.72	0.82		0.91,0.97,1.06
19V	8.8	115.29	175.11	57.57	35.22	20.05,22.43		5.33	2.11			0.70,0.82
20V	8.77	118.43	174.22	60.89	35.97	22.01,22.03		5.42	1.94			0.85,1.06
21A	9.66	129.42	174.95	50.01	25.82			5.77	1.22			
22E	9.08	120.44	174.98	54.36	27.26	34.59		5.07	1.85,1.99			2.03,2.23
23K	8.57	125.98	175.46	58.19	34.28			4.43	1.77,2.15	1.55,1.62		
24E	8.71	130.27	175.24	55.82	31.27	36.13		4.56	1.78,1.96			2.15,2.22
25G	8.77	116.46	174.8	47.2				3.66,4.03				
26D	8.92	127.16	174.88	55.38	41.61			4.67	2.45,2.89			
27D	7.86	119.37	175.12	53.35	44.9			5.13	2.67,2.76			
28T	9.67	117.59	174.08	60.97	68.41	24.76		4.68	3.6			0.73,1.22
29I	8.61	120.6	176.43	59.31	39.53	18.27,26.11		3.88	1.67	0.65		0.73,1.27,1.52
30L	8.07	125.23	176.91	56.18	42.85	26.89	23.34,25.99	4.3	1.89,2.05	0.43,0.66		1.86

Residue	H	N	C	Ca	Cb	Cg	Cd	Ha	Hb	Hd	He	Hg
31I	9.02	115.99	174.68	61.19	40.37	26.71		4.65	2.28	0.34		0.38,1.03,1.10
32Y	7.74	122.98	174.05	59.01				4.58				
33V	7.3	120.08	172.94	59.04	35.76	19.71		4.69	2.03			0.70,0.81
34N	7.79	119.6	175.88	50.3	40.06			5.02	3.54,3.85	6.60,7.89		
35A	8.27	123.48	180.41	54.57	17.4			4.62	1.48			
36A	8.88	120.41	180.97	55.09	17.79			4.22	1.46			
37F	8.22	121.26	177.32	62.41	40.16			4.28	3.05,3.32	7.12	6.89	
38E	7.66	120.63	179.42	60.04	30.42	37.62		4.22	2.13,2.23			2.36,2.57
39Y	8.36	119.36	178.25	60.28	37.93			4.21	3.10,3.10		6.83	
40L	8.19	119.82	176.53	57.58	43.04	25.95	21.73,24.09	3.79	1.20,1.75	0.56,0.77		0.88
41T	7.13	101.38	174.57	63.09	70.99	22.13		3.95	4.45			0.57,0.66
42G	8.11	109.55	173.42	45.72				3.61,4.25				
43Y	7.72	119.1	174.55	57.86	41.04			4.5	2.74,2.98	7.07	6.88	
44S	8.88	114.15	177.54	57.26	65.54			4.79	3.96,4.14			7.79
45R	9.81	123.45	176.74	60.37	28.88		43.58	3.72	2.04,2.11	3.19,3.28		1.66,1.74
46D	8.22	116.07	177.88	56.89	40.56			4.37	2.52,2.70			
47E	7.79	117.71	176.34	57.3	31.6			4.17	2.10,2.32			2.54,2.69
48I	7.36	116.52	174.75	63.41	41.18	19.78		3.46	1.53	0.67		0.70,0.83,0.85
49L	7.45	113.83	177.5	56.42	42.04		23.34,25.65	3.39	1.29,1.40	0.39,0.63		1.23
50Y	8.94	111.59	174.33	58.55	33.98			3.93	3.28,3.38	6.95	6.9	
51Q	7.66	115.2	174.69	53.92				4.75	1.90,2.11			2.22,2.30
52D	8.46	122.4	177.1	54.94	44.93			5.04	2.31,2.76			
53C	9.04	128.47	172.17	58.35	34.04			3.51	3.09,3.39			
54R	9.82	115.46	178.68	58.28			43.55	4.25	2.05,2.15	3.43,3.47	8.18	
55F	7.67	123.89	177.68	59.95	37.11			4.49	3.17,3.22	7.18		
56L	7.3	123.65	177.71	56.88	42.43		25.15,26.93	3.93	1.06,1.50	0.67,0.76		1.03
57Q	7.82	116.61	176.13	57.18	28.11	33.47		4.25	1.85,2.17		6.72	2.52,2.69

Residue	H	N	C	Ca	Cb	Cg	Cd	Ha	Hb	Hd	He	Hg
58G	7.68	105.01	174.65	47.16				3.59,3.72				
59D	8.24	124.64	176.03	53.86	40.94			4.71	2.50,2.74			
60D	8.46	124.98	176.55	52.56	40.19			4.83	2.51,3.59			
61R	7.66	120.08	176.68	55.71	31.02	26.39	42.05	4.57	1.73,1.87	3.23,3.31		1.29,1.38
62D	8.98	124.57	175.55	53.25	39.86			5.06	2.52,2.91			
63Q	7.29	118.32	178.11	56.21	31.1	32.25		4.63	1.88,1.90		5.70,6.38	2.11,2.36
64L			179.35	58.19				4.17				
65G	9.05	104.07	175.36	47.31				3.71,3.71				
66R	7.42	118.47	177.75	60.73	31.19		43.77	3.74	1.83,2.27	3.24,3.32		1.66,1.73
67A	7.44	119.05	180.77	55.16	18.57			4.03	1.55			
68R	7.14	116.75	179.5	60.08	33.26		43.66	4.02	1.77,1.91	3.20,3.35		1.47,1.60
69I	8.06	121.93	177.4	65.11	39.06	18.21		3.95	1.73	0.54		0.70,1.40,1.44
70R	8.12	117.63	179.59	60.24	30.71			4.01	1.81,1.98	3.58,3.61		1.44,1.53
71K	7.78	119.21	177.56	59.37	33.06	25.07		4.08	1.83,1.89	1.51,1.60	3.00,3.00	1.35,1.47
72A	7.64	121.13	180.22	55.95	17.91			4.02	1.51			
73M	8.59	115.79	179.75	60.39	36.25	31.92		4.04	2.09,2.20			2.66,2.74
74A	8.03	121.73	179.94	54.99	18.42			4.19	1.6			
75E	7.88	115.62	176.72	56.29	30.51	36.83		4.3	1.96,2.08			2.22,2.33
76G	8.02	109.32	173.97	47.16				3.63,4.07				
77R	8.48	120.49	172.21	52.81	33.22	27.57	43.36	5	1.91,2.08	3.29,3.29		1.51,1.68
78P			176.03	62.19	33.06	27.56	52.18	5.18	2.09,2.51	3.65,3.87		1.43,1.90
79C	8.7	113.42	171.82	56.1	31.98			4.92	2.65,3.01			1.51
80R	8.52	124.55	175.1	55.6			43.75	5.67	1.81,1.89	3.24		1.65,1.71
81E	8.58	124.6	173.71	53.94	29.47	36.35		4.75	1.90,1.90			2.11,2.11
82V	8.57	121.91	175.69	61.1	33.42	21.9		5.11	1.89			0.81,0.87
83L	8.88	126.09	175.4	52.96	46.98	27.23	24.83,27.07	4.83	1.28,1.38	0.51,0.61		1.2
84R	9.91	124.28	176.47	56.38				3.72	2.45	3.28		1.35,1.38

Residue	H	N	C	Ca	Cb	Cg	Cd	Ha	Hb	Hd	He	Hg
85N	8.65	129.31	170.72	50.04	46.32			5.01	2.17			
86Y	8.43	115.93	176.5	57.25	40.86			5.07	2.53,2.73	7.03	6.75	
87R	9.03	120.8	178.54	55.92	32.12		43.8	4.9	1.71,1.88	2.74,2.89		1.23,1.35
88K	7.71	124.07	177.15	59.77	29.95	25.45		3.6	1.37,1.71	1.21,1.21	2.75,2.86	1.11,1.11
89D	7.79	113.63	176.98	53.91	39.84			4.32	2.56,3.06			
90G	8.35	109.64	174.73	45.11				3.76,4.40				
91S	8.38	117.64	172.8	59.24	63.92			4.27	3.88,3.96			
92A	8.24	124.59	177.64	50.61	21.53			4.9	0.98			
93F	8.8	118.26	171.91	55.84	35.23			4.69	2.09,2.23			
94W	8.07	122	172.73	56.23	30.03			4.62	2.73,2.85	7.29	7.25,10.40	
95N	9.12	126.57	174.61	50.5	41.74			5.32	2.61,2.63	6.77,7.45		
96E	9.52	126.58	173.5	56.07		37.03		4.6	1.84,1.90			2.21
97L	9.1	129.81	175.58	54.2	42.63		23.05,24.97	4.5	0.86,1.10	0.38,0.51		0.71
98S	8.41	117.36	175.65	56.16	63.02			5.34	3.59,3.59			5.67
99I	8.72	127.99	176.13	60.63	41.67		18.48	5.03	1.66	0.73		1.09,1.34,1.52
100T	9.18	121.86	172.52	60.69	71.89	20.33		5.02	4.17			1.23,1.43
101P			175.97	61.65	32.29			4.66	1.85,2.10			1.66,1.72
102V	9.09	123.91	174.2	62.23	36.19	20.91,21.97		4.06	1.72			0.79,0.87
103K	8.5	127.65	176.31	55.7	32.84	24.87		4.65	1.77	1.64	2.94	1.27,1.27
104S	8.71	120.32	175.76	57.14	64.36			4.58	3.90,3.94			
105D	8.74	126.2	177.02	55.9	40.96			4.52	2.50,2.62			
106F	8.21	118.13	176.16	59.51	39.22			4.5	3.07,3.09	7.25		
107D	7.58	116.14	175.74	53.26	41.58			4.57	2.55,2.92			
108Q	7.98	113.62	175.42	57.2	26.98			4.11	2.25,2.32		.78	2.54,2.93
109R	8.1	120.03	175.17	55.93	None		43.58	4.43	1.55,1.78	3.17		1.40,1.46
110T	8.24	117.35	172.21	63.21	69.77	22.06		4.51	3.79			0.73,0.73
111Y	8.53	125.33	173.8	56.34	43.94			5.28	2.64,3.06	6.93	6.82	

Residue	H	N	C	Ca	Cb	Cg	Cd	Ha	Hb	Hd	He	Hg
112F	9.09	114.94	174.81	56.04	44.99			5.51	2.83,2.83	6.82	6.94	
113I	8.7	120.58	173.48	57.92	33.13	14.60,24.78	13.26	4.43	1.66	0.73		1.23,1.56
114G	9.35	116.79	171.68	44.55				3.45,3.45				
115I	8.84	126.75	175.98	60.74	39.12	29.36		5.01	1.9			1.24,1.36
116Q	8.99	127.64	174.92	56.67	27.35			3.41	1.86,2.01			2.11,2.36
117K	8.2	120.02	173.69	55.6	36.98	24.62		4.61	1.85,1.91	1.56	2.89,3.01	1.37,1.46
118D	8.7	121.79	177.79	54.03	40.54			5.37	2.43,3.03			
119V	9.23	123.16	176.4	60.37	31.28	19.92,22.23		5.03	2.64			0.86,0.89
120S	9.22	119.75	175.49	64.31				4.45	3.90,4.05			
121R	8.58	119.36	177.98	59.21			43.3	4.06	1.78,1.82	3.14,3.28		1.58,1.66
122Q	7.76	119.77	178.24	59.59				3.87	1.86,1.93			2.00,2.12
123V	7.94	120.35	178.66	66.32	31.96	22.49		3.36	1.43			0.46,0.46
124E	7.88	119.33	179.51	59.36	29.64			4.03	1.72,1.87			1.98,2.13
125L	7.88	121.47	178.54	58.61	41.52	26.27	23.82	4.11	1.46,1.86	0.77,0.78		0.93
126E	7.92	119.24	179.49	60.3	29.29	37.23		3.79	2.04,2.11			2.43
127R	8.06	120.89	178.63	59.41		27.34	43.1	4.02	2.43,2.43	3.13,3.17		1.51,1.58
128E	8.1	121.93	179.1	59.61	29.58	36.36	None	4.01	1.98,2.14			2.43
129L	8.31	119.46	177.81	58.14	42.09	27.33	24.03	3.98	1.91,2.13	0.85,0.93		1.48
130A	7.88	120	180.88	55.34	18.55			3.99	1.51			
131E	7.83	117.71	179.15	59.14	29.78	36.55		4.06	1.77,1.87			2.15,2.26
132L	8.07	119.9	179.79	57.48	42.84		23.64	4.14	1.69,1.88	0.87,0.88		1.53
133R	8.25	117.97	176.78	58.19	31.27	27.89	44.28	4.06	1.80,1.87	3.06,3.09		1.52,1.63
134A	7.4	120.57	177.31	52.59	19.44			4.32	1.46			
135R	7.61	120.39	174.13	54.64		27.02	43.98	4.57	1.88,2.03	3.23,3.27		1.77,1.84
136P			176.67	63.15	32.26		50.73	4.45	2.03,2.31	3.64,3.85		1.75,1.88
137K	8.4	122.6	174.94	54.38	32.58			4.6	1.75,1.88	1.53,1.67		
138P			176.59	63.62	32.25	27.55		4.4	2.31,2.31	4.01,4.01		1.94,2.03

Residue	H	N	C	Ca	Cb	Cg	Cd	Ha	Hb	Hd	He	Hg
139D	8.32	119.4	176.21	54.41	41.2			4.56	2.64,2.70			
140E	8.21	121.33	176.22	56.78	30.55	36.54		4.3	1.96,2.08			2.26
141R	8.24	122.03	174.92	56.11	31.1			4.34	1.78,1.90	3.28		1.60,1.65
142A	7.86	130.91	182.42	53.94	20.46			4.12	1.34			

Declaration

I hereby declare that this dissertation is produced autonomously and independently and without using any unauthorized help. This dissertation has not been handed in another institution or university. I do not have unsuccessful attempts of obtaining a doctoral degree.

Jülich, 12.11.2012

Raj Rani