

Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Chemie und Pharmazie  
der Ludwig-Maximilians-Universität München

**Studies on Structure and Dynamics of Components of  
Phytochrome Mediated Light Signaling  
in Plants and Cyanobacteria**

**Christian Benda**

**aus**

**München**

2002



## Erklärung

Diese Dissertation wurde im Sinne von §13 Abs. 3 bzw. 4 der Promotionsordnung vom 29. Januar 1998 von Prof. Dr. D. Oesterhelt und von apl. Prof. Dr. W.-Gärtner betreut.

## Ehrenwörtliche Versicherung

Diese Dissertation wurde selbständig, ohne unerlaubte Hilfe erarbeitet.

München, am 05.06.2002

---

Christian G. Benda

Dissertation eingereicht am: 30.08.2002

1. Berichterstatter: Hon.-Prof. Dr. Dieter Oesterhelt

2. Berichterstatter apl.-Prof. Dr. Wolfgang Gärtner

Tag der mündlichen Prüfung: 12.02.2003



Meinen Eltern



# Contents

<b>Contents</b>	<b>1</b>
<b>1 Summary</b>	<b>1</b>
<b>2 Introduction</b>	<b>3</b>
2.1 Light and Photosensory Perception	3
2.1.1 Phytochrome - a ubiquitous sensory pigment in plants and bacteria	5
2.1.2 Spectroscopic properties of phytochromes	7
2.1.3 Bacterial phytochromes	8
2.1.4 Phytochrome signaling	10
2.2 Conceptual Formulation	15
<b>3 Materials and Methods</b>	<b>17</b>
3.1 Materials	17
3.1.1 Equipment	17
3.1.2 Bacterial and yeast strains	19
3.1.3 Plasmids	20
3.2 General Molecular Biological Techniques	21
3.2.1 Culture of <i>E. coli</i> DH5 $\alpha$ cells for plasmid growth	21
3.2.2 Purification of plasmid DNA from <i>E. coli</i> cells	21
3.2.3 Isolation of total RNA from <i>Arabidopsis thaliana</i>	21
3.2.4 General PCR protocol for DNA amplification	21
3.2.5 Reverse transcription polymerase chain reaction (RT-PCR)	22
3.2.6 Purification of DNA after enzymatic reactions	22
3.2.7 Preparation of vector DNA for ligation reactions	22
3.2.8 Preparation of PCR-derived DNA fragments for ligation reactions	23
3.2.9 Ligation of DNA molecules	23
3.2.10 Restriction analysis of plasmid DNA	23
3.2.11 Analysis of DNA by agarose gel electrophoresis	23
3.2.12 Sequencing of plasmid DNA	24
3.2.13 Introduction of point mutations by site-directed mutagenesis	24

3.2.14	Construction of mutant genes by PCR	24
3.2.15	Isolation of DNA fragments from agarose gels	25
3.2.16	Determination of DNA concentration in aqueous solution	25
3.2.17	Preparation of electrocompetent <i>E. coli</i> cells	25
3.2.18	Transformation of <i>E. coli</i> cells	25
3.2.19	Cloning of <i>PIF3</i> from <i>A. thaliana</i>	26
3.2.20	Cloning of <i>PHYA</i> and <i>PHYB</i> from <i>A. thaliana</i> into the vector pPICZ A	27
<b>3.3</b>	<b>Protein Chemical Methods</b>	<b>29</b>
3.3.1	Test for pET28a(+) directed gene expression	29
3.3.2	Standard SDS-polyacrylamide gel electrophoresis (SDS-PAGE)	29
3.3.2.1	IEF gel electrophoresis	30
3.3.3	Immuno-detection of immobilized proteins (Western Blotting)	30
3.3.4	Heterologous expression of proteins in <i>E. coli</i>	31
3.3.5	Heterologous expression of RcpA and RcpB	31
3.3.6	Heterologous expression of L-selenomethionine labeled RcpB	31
3.3.7	Purification of overexpressed proteins from <i>E. coli</i>	33
3.3.7.1	Purification of wild type RcpA and RcpB	33
3.3.8	Purification of overexpressed proteins from <i>P. pastoris</i> and <i>H. polymorpha</i>	34
3.3.9	Estimation of protein apparent molecular weights by gel filtration	34
3.3.10	<i>In vitro</i> Reconstitution of heterologously expressed phytochromes	35
3.3.11	Isolation of PIF3 inclusion bodies from <i>E. coli</i> cells	36
3.3.12	Refolding of proteins from inclusion bodies	36
3.3.13	Thrombine cleavage	36
3.3.14	Phosphorylation assays	36
<b>3.4</b>	<b>Cloning and Expression in Yeast</b>	<b>38</b>
3.4.1	Preparation of linearized plasmid DNA for yeast transformation	38
3.4.2	Preparation of electrocompetent <i>P. pastoris</i> cells	38
3.4.3	Transformation of <i>P. pastoris</i> cells	39
3.4.4	Control of gene expression in recombinant <i>P. pastoris</i> strains	39
3.4.5	Fermentation of <i>P. pastoris</i> and <i>H. polymorpha</i>	40
3.4.5.1	Fermentation of yeasts for expression of <sup>15</sup> N-labeled proteins	42
<b>3.5</b>	<b>Spectroscopic Methods</b>	<b>44</b>
3.5.1	UV/VIS Spectroscopy of recombinant phytochromes	44
3.5.2	Circular dichroism spectroscopy	44
3.5.3	<sup>1</sup> H- <sup>15</sup> N HSQC Nuclear Magnetic Resonance Spectroscopy	46
3.5.3.1	<sup>1</sup> H- <sup>15</sup> N HSQC spectra of 59 kD phytochrome A	46
<b>3.6</b>	<b>Protein Crystallographic Methods</b>	<b>47</b>

3.6.1	Protein crystallization techniques	47
3.6.2	Crystallization of RcpA	48
3.6.3	Crystallization of RcpB and SeMet-RcpB	48
3.6.4	X-Ray structural analysis	48
3.6.5	Data collection	48
3.6.5.1	RcpA native data collection	49
3.6.5.2	RcpB data collection	49
3.6.6	Data processing and reduction	50
3.6.7	Estimation of the number of molecules per unit cell	50
3.6.8	Solution of the Phase Problem	51
3.6.8.1	Patterson search	51
3.6.8.2	Solution of the phase problem for RcpA by MR	53
3.6.8.3	The Multiple Isomorphous Replacement method (MIR)	53
3.6.8.4	The Multiple Anomalous Wavelength Dispersion Method (MAD)	54
3.6.8.5	Solution of the phase problem for RcpB by MAD	55
3.6.9	Model Building and Refinement	55
3.6.9.1	Refinement of the structures of RcpA and RcpB	57
3.6.10	Validation of Molecular Structures	57
3.6.11	Visual Representation of Molecular Structures	58
<b>4</b>	<b>Results</b>	<b>59</b>
4.1	X-ray Structural Analysis of RcpB	59
4.1.1	Expression and purification of wild type RcpB	59
4.1.2	Expression and purification of SeMet-RcpB	60
4.1.3	Crystallization and diffraction data collection of RcpB	62
4.1.4	Determination of the phase angles for RcpB by MAD	65
4.1.5	Model building and structure refinement of RcpB	66
4.2	X-ray Structural Analysis of RcpA	69
4.2.1	Expression and purification of recombinant RcpA	69
4.2.2	Crystallization and diffraction data collection of RcpA	70
4.2.3	Solution of the phase problem for RcpA by Patterson search	72
4.2.4	Structure refinement of RcpA	72
4.3	Crystal Structures of RcpA and RcpB	76
4.3.1	Overall structure of phospho-RcpA and apo-RcpB.	76
4.3.2	Intermolecular interaction in the homodimers of RcpA and RcpB.	78
4.3.3	Comparison with other receiver domains	83
4.3.4	Active site architecture	85

4.3.4.1	Phospho-RcpA	85
4.3.4.2	Apo-RcpB	87
4.3.5	Conformational differences between phospho-RcpA and apo-RcpB	87
4.3.6	Comparison of RcpA and B with other receiver domains	90
<b>4.4</b>	<b>The N-Terminal Receptor Domain of Plant Phytochrome A</b>	<b>91</b>
4.4.1	Expression and purification of 59 and 65 kD PhyA	91
4.4.1.1	Expression, <i>in vitro</i> reconstitution and purification of 59 / 65 kD PhyA	92
4.4.1.2	UV/VIS Spectroscopic characterization	94
4.4.1.3	Mass spectrometric analysis	96
4.4.1.4	Isoelectric focussing	96
4.4.2	Circular dichroism	98
4.4.3	NMR spectroscopic characterization of recombinant phytochromes	100
4.4.4	Crystallization screening	104
4.5	Cloning and expression of phytochrome interacting factor 3	105
4.6	Cloning of <i>A. thaliana</i> PHYA, PHYB and PIF3	106
<b>5</b>	<b>Discussion</b>	<b>108</b>
5.1	Crystal Structures of RcpA and RcpB	108
5.1.1	Function and activation of the chemotaxis receiver domain CheY	108
5.1.2	Dimerization of RcpA and RcpB	111
5.1.3	Activation of the cyanobacterial response regulators RcpA and RcpB	115
5.1.3.1	The active and inactive state in RcpA and RcpB	115
5.1.3.2	Dynamic or static?	119
5.1.3.3	Conclusion	120
5.2	59 and 65 kD N-terminal fragments of oat phytochrome A	121
5.2.1	Expression and purification	121
5.2.2	Photoactivity of phy fragments	122
5.2.3	Crystallization trials and NMR-spectroscopy	123
<b>6</b>	<b>Bibliography</b>	<b>127</b>
<b>7</b>	<b>Appendix</b>	<b>138</b>
	Sparse Matrix Screening	138
	Abbreviations	142

**Danksagung**

**144**

***Curriculum Vitae***

**146**

## 1 Summary

From a structural point of view, phytochrome mediated light signaling in plants and bacteria constitutes a scarcely understood system. Albeit the tremendous research activity in the field, no X-ray or NMR-based structural information is available. For the first time now, the crystal structures of two cyanobacterial response regulators (RR), which directly interact with bacterial phytochrome receptors, have been solved. This could be accomplished by heterologous expression, purification and crystallization of the proteins, and X-ray structure determination, employing multiple anomalous wavelength dispersion (MAD) and molecular replacement (MR) techniques.

The two RRs, **RcpA** and **RcpB**, belong to the family of CheY-homologous receiver modules and possess a similar three dimensional folding topology of alternating  $\beta$ -sheets and  $\alpha$ -helices  $[\beta/\alpha]_5$ , that form a central, five-stranded, parallel  $\beta$ -pleated sheet, surrounded by five  $\alpha$ -helices. In the crystal structures, as well as in solution, the receivers were found to form stable homodimers. This dimerization is mediated by specific interactions of a surface area, formed by helix 4,  $\beta$ -strand 5 and helix 5, which represents a new feature of this class of proteins. Comparison of both interface areas of RcpA and RcpB, revealed a conserved structural motif for dimerization that, by means of database mining, could also be found in several other RRs. This motif was found in all RRs associated with a phytochrome-like receptor histidine kinase. It was therefore postulated, that specific dimerization, via the aforementioned interface, represents a highly conserved characteristic of bacterial phytochrome RRs, and allows to classify them as a new subgroup of the CheY superfamily.

Response regulators are aspartate kinases, that receive information by a phosphotransfer event. RcpA crystallized in the phosphorylated and RcpB in the unphosphorylated [apo-] form, a circumstance, which permitted the structural comparison of the putative active state (RcpA), with the inactive state (RcpB) and the formulation of a possible mechanism through which phytochrome RRs might function. According to this, the RRs are monomeric in the inactive apo-form, and as monomers, they interact with the cognate phytochrome receptor. Upon light absorption by the photoreceptor, they become phosphorylated, which reduces their affinity to the histidine kinase domain of the phytochrome and promotes homodimerization. The homodimeric form interacts with downstream components and subsequent dephosphorylation (either spontaneous or through a specific phosphatase) initiates the disintegration of the complex.

Two fragments of a plant phytochrome, PhyA from oat (*Avena sativa*), were subject of the second part of this study. The fragments comprised the 59 and 65 kD N-terminal receptor domains of PhyA and differed in the very N-terminal part (65 residues), which is believed to form a helical sub-domain involved in the interaction with downstream components.

The two receptor fragments were heterologously expressed in yeast and purified to homogeneity. *In vitro* reconstitution with the chromophore derivative phycocyanobilin (PCB) yielded fully photoreversible holo-proteins, which exhibited the characteristic UV/VIS spectral properties, comparable to native full-length phytochrome. The fragments were characterized and subjected to crystallization trials.

Circular dichroic (CD) methods showed, that the two constructs differ in their behavior upon light absorption. In 65 kD PhyA, the transformation from the red light- (Pr) to the far-red light- (Pfr) absorbing form is accompanied by a reverse refolding, that could be monitored by CD spectroscopy. In the truncated 59 kD fragment, this refolding was not observed. It was therefore stated, that the N-terminal peptide must be responsible for the measured CD effect, and that signaling involves the refolding (unfolding) of this mostly helical domain.

In order to monitor light induced dynamics, the phytochrome fragments were labeled with the nitrogen-15 isotope to perform  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear magnetic resonance spectroscopy (HSQC). For this purpose, a protocol for the large scale expression of nitrogen-15 labeled phytochrome in yeast was established. The results from  $^1\text{H}$  NMR and  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR experiments showed that the fragments were not specifically folded, but rather represent dynamical entities with high conformational flexibility. In the context of results obtained from reconstitution experiments and UV/VIS spectroscopy this finding was particularly interesting. Obviously, full establishment of autocatalytic chromophore incorporation and the formation of photoreversibility and the characteristic spectroscopic profile does not necessarily involve the complete specific and/or stable folding of the receptor domains.

## 2 Introduction

### 2.1 Light and Photosensory Perception

Light is the most fundamental source of energy on our planet, and life would never have evolved without it. The sun existed long before the first life-form arose and ever since, it has issued its power and submerged earth into an everflowing stream of electromagnetic energy. Through billions of years of evolution the harmonically oscillating illumination of the biosphere has exerted its influence on the unfolding of biological diversity. As a consequence, every living organism, whether directly or indirectly, depends on the continual flow of light energy. A great multitude of organisms has dedicated its entire existence to the deployment and transformation of light in order to convert it into a more versatile form of chemical energy. It is thus not surprising, that evolution has equipped many life-forms with the ability to perceive, analyze and interact with light in all its diverse manifestations.

From a physical point of view, electromagnetic radiation interacts with particles in several ways, depending on the energy of the photons. Solid matter for example, appears colored if the electronic energy levels in the molecular building blocks coincide with the energy of photons of the visible spectrum. As a result of the interaction, the spectral composition of the incident light is altered which is experienced as the sensation of color. It is with the aid of particular organic molecules, also referred to as chromophores, that nature has created protein-based macromolecular systems for the specific interaction with light. These biological chromophores engaged in light perception can be divided into three classes: tetrapyrroles (heme, bilins), polyenes (retinal) and aromatics (flavins, p-hydroxy cinnamic acid). Embedded in a proteinaceous environment, they are utilized to form biological light receptors, that constitute an organism's link between the world of photophysical information and the fine-tuning of physiological processes as a response. Most important among them are the visual pigments, light sensitive molecules belonging to the rhodopsin family, which provide vertebrates as well as invertebrates with the ability of vision. These molecules consist of a seven transmembrane helical protein moiety of approx. 40 kD and a covalently attached retinal or retinal derivative as a chromophore. Besides the visual pigments, entirely dedicated to the animal kingdom, there is a number of photoreceptors which are involved in the so called nonvisual light perception in animals, plants and bacteria. Among these, and certainly best known, is a class of receptors that is exclusively found in the kingdom of plants and some bacteria, the so called phytochromes (phy) [Kendrick and Kronenberg, 1994]. As described in more detail in the following paragraphs, they constitute a family

of sensory photoreceptors, which allow plants and bacteria to monitor primarily the red/far-red spectral region and respond to changes in the surrounding environment. Other nonvisual photoreceptors are specialized in the perception of blue and ultraviolet light, like the flavin-containing phototropins (phot), engaged in movement events of plants (phototropism) (Lin, 2002), and the family of cryptochromes (cry), receptors that entrain the endogenous circadian clocks and other processes, related to photoperiodicity in plants and animals (Cashmore, et al., 1999). Sensory rhodopsin I (SR-I) from *Halobacterium salinarum* and photoactive yellow protein (PYP) from *Ectothiorhodospira halophila* are photosensors involved in phototactic responses in these archaea (Hellingwerf, et al., 1996, Krah, et al., 1994, Marwan, et al., 1995).



Figure: Botanical illustration of the model plant *Arabidopsis thaliana*.

### 2.1.1 Phytochrome - a ubiquitous sensory pigment in plants and bacteria

Phytochromes constitute one of the best characterized non-visual photoreceptor system. They are ubiquitous in higher and lower plants and also occur in several algae and bacteria. Plants, more than motile organisms, need to analyze and adapt to their surrounding environmental conditions in order to optimize growth and development. Unable to simply change their location by moving in order to maintain optimal conditions, plants deploy a group of sensory receptors, enabling them to fine-tune and regulate their biochemical apparatus in response to the everchanging environment.

As the ultimate source of energy, light doubtlessly constitutes the most important and most influential environmental parameter for plants, and they have thus learned to exhaustively analyze light conditions in a given habitat. By means of a complex system of different photoreceptors, which detect a wide spectral range, plants are able to monitor the presence, absence, spectral quality, fluence rate, directional and diurnal duration of the incident light signals. This permits modulation of germination, growth rates, detection of neighboring plants, induction of flowering and all aspects generally subsumed under the term *photomorphogenesis* (Chory, 1997, Chory, et al., 1996, Kendrick and Kronenberg, 1994, Millar, et al., 1994, Quail, 1994). Photosensory proteins from plants include the blue and UV-A light receptors phototropin, responsible for phototropism (the bending/movement response in accordance to the direction of light) and the cryptochromes, engaged in the entrainment of the circadian clock. The largest group is formed by the red (R) and far-red light (FR) photoreceptors named phytochromes. Their first representative was described more than forty years ago (Buttler, et al., 1959). Ever since, our knowledge of these important plant photoreceptors has significantly grown by virtue of considerable efforts that have been put into the field. In the model plant *Arabidopsis thaliana*, which has been granted most of these research activities (illustration on previous page), the phytochromes are encoded by a gene family consisting of five members (PHY A-E) (Quail, et al., 1995).

Phytochromes of all plants so far investigated share some common characteristics. They are soluble, homodimeric chromoproteins with a molecular weight of approx. 125 kD per subunit, each of which folds into two distinct structural domains that are linked by a so called hinge region. The amino-terminal domain is the photosensory part of the protein and responsible for the initial interaction with incident photons. Within this domain lies a highly conserved region that autocatalytically attaches a chromophore molecule which becomes covalently bound to a conserved cysteine residue *via* a thioether linkage. The universal chromophore in plant phytochromes is the open chain tetrapyrrole phytochromobilin (PΦB), depicted in Fig. 2.1. The C-terminal domain of the



to E conformation as indicated in Fig. 2.1. This internal structural change entails a yet unknown conformational change of the protein backbone and shifts the receptor to the biologically active far-red absorbing form. As a consequence, an intracellular signal transduction process is triggered, that finally alters the transcription of selected genes and thereby controls photomorphogenesis to match the given environmental conditions. In turn, the absorption of far-red light by the activated  $P_{fr}$  form immediately regenerates the inactive  $P_r$  form, which counteracts to the  $P_{fr}$  initiated processes and reverts the signal.

### 2.1.2 Spectroscopic properties of phytochromes

Being photochromic molecules, phytochromes exhibit a typical action spectrum which can be used for a preliminary characterization. As outlined above, these soluble, light driven switches exist in two stable, photoconvertible forms, the  $P_r$  form and the  $P_{fr}$  form. Physically, they can be discriminated by their different absorption behavior in the visible region of the electromagnetic spectrum. Figure 2.2 shows a typical UV/VIS spectrum of a recombinant phytochrome.

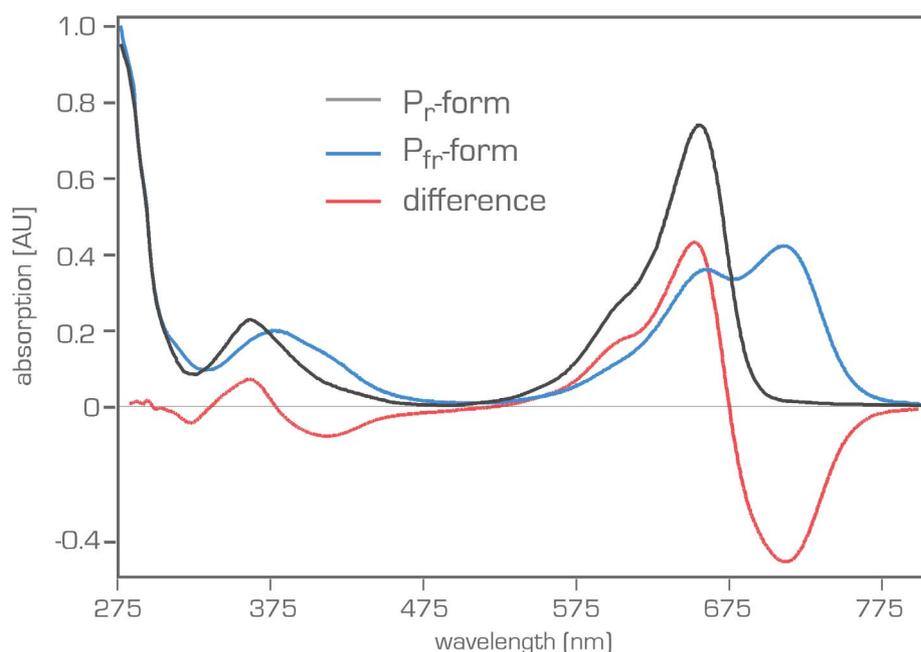


Figure 2.2. **Spectral Properties of phytochromes.** A typical UV/VIS spectroscopic recording of a recombinant holo-phytochrome A from oat (65 kD N-terminal domain) assembled with the chromophore phycocyanobilin (PCB) is shown. The inactive red light absorbing  $P_r$  form has a maximum absorbance around 665 nm (marked black). Irradiation with red light triggers the photoconversion to the active far-red absorbing  $P_{fr}$  form with an absorption maximum at 730 nm (marked blue). Due to the overlapping spectra of the two isomers, a photoequilibrium exists dependent on the light quality. To estimate the exact

amount of photoactive protein, difference spectra (red) are calculated to determine the position of the absorption maxima and the differential change in absorption ( $\Delta A$ ) between the two forms upon photoconversion.

The delocalized  $\pi$ -electron system in the incorporated tetrapyrrole chromophore P $\Phi$ B, which extends over all four pyrrole rings A-D (Fig. 2.1), is responsible for the absorption of light in the visible region. All phytochromes possess a prominent red absorption band in the P<sub>r</sub> state with an absorption maximum around 665 nm. Upon light absorption and photoconversion to the P<sub>fr</sub> form, this absorption maximum experiences a bathochromic shift with a resulting maximum in the far-red region, around 730 nm. Additionally, and apart from the absorption band at 280 nm originating from aromatic side chains in the protein moiety, there is a comparably weak blue absorption band with maxima at approx. 380 and 400 nm for the P<sub>r</sub> and P<sub>fr</sub> form, respectively.

Even though the two photoactive states have distinct absorption maxima, their absorption bands overlap, which leads to the formation of a photoequilibrium in dependence to the wavelength. Optimally the P<sub>r</sub> state forms to 100 % *in vitro*, whereas the P<sub>fr</sub> state only reaches levels around 70%. By calculating a P<sub>r</sub> - P<sub>fr</sub> difference spectrum, the content of photoactive phytochrome in a given preparation can be estimated, if the extinction coefficient is known (Lambert-Beer's Law). As a measure of the purity of a phytochrome preparation, the ratio between the chromoprotein P<sub>r</sub> absorption ( $\lambda_{\text{max}} = 665 \text{ nm}$ ) and the protein band ( $\lambda_{\text{max}} = 280 \text{ nm}$ ) can be evaluated and is referred to as the specific absorption ratio SAR.

Although immense research activity has been focused on the mechanism underlying phytochrome action since their first discovery in plants, the exact function remains unknown. In most plants, up to five phytochromes are present (PhyA-E in *A. thaliana*) and their physiological function is only partly differential and overlap in some aspects (Devlin, et al., 1998, Devlin, et al., 1999, Quail, 1998, Smith, 2000, Whitelam, et al., 1998). Each of them interacts with several downstream factors some of which have been identified for PhyA and PhyB but they also seem to share common signaling pathways among each other and even with other photoreceptors (cry and phot) (Quail, 2002).

### 2.1.3 Bacterial phytochromes

So far, phytochromes of plants or any other organism have successfully avoided to divulge their atomic architecture, although several groups have directed their efforts towards structure determination over many years. It has thus been a tremendous progress in the field, when recently genes encoding proteins homologous to plant phytochromes were found in several bacteria. These bacterial phytochromes or

bacteriophytochromes [BphPs] (Vierstra and Davis, 2000) have a phytochrome-like N-terminal sensor domain that is capable of tetrapyrrole binding and formation of a photoactive holo-protein, which spectroscopically resembles those of plant homologues. Most of these bacterial phytochromes possess a C-terminal domain that is highly homologous to histidine kinase domains [HKD] known from the classic bacterial two-component signaling system. Clearly, this constitutes a major step towards an understanding of phytochrome functioning since the two-component system is one of the best investigated and understood signaling systems nowadays and it allows to approach the principles relating light sensing with signal and response generation from a prokaryotic starting point as a model system. Figure 2.3 gives an overview of the general domain organization of this receptor class.

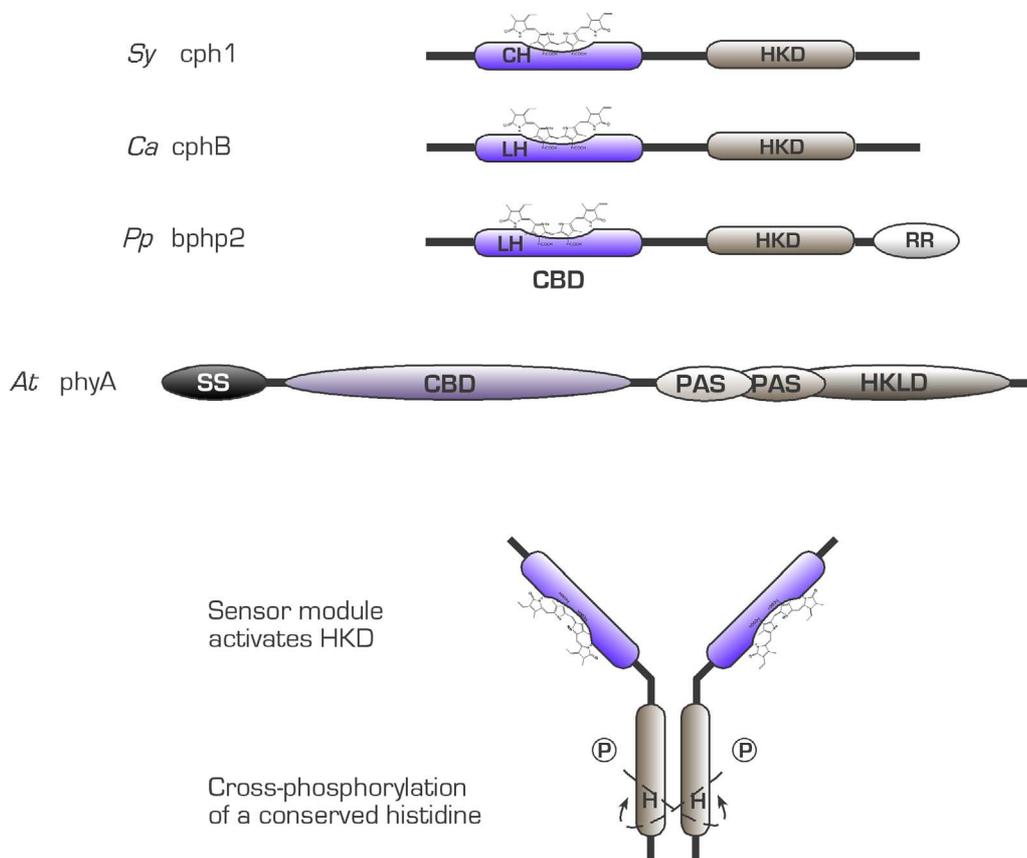


Figure 2.3. **Domain organization in bacteriophytochromes.** Shown are the domain organization of the phytochrome homologous proteins *cph1* from *Synechocystis* sp. [*Sy*], *cphB* from *Calothrix* sp. [*Ca*] and *bphp2* from *P. putida* [*Pp*] in comparison to a higher plant phytochrome, PhyB from *A. thaliana* [*At*]. CBD, chromophore binding domain; HKD, histidine kinase domain; RR, response regulator domain; PAS, Per-Arndt-Sim repeats; HKLD, histidine kinase like domain; SS, serine rich domain. The chromophore attachment site is indicated by the conserved pair of amino acids [CH and LH]. Two-component histidine kinases are typically assembled as homodimers (bottom) and act by cross phosphorylation of a conserved histidine residue upon receiving an environmental light stimulus.

Bacteriophytochromes were first discovered in 1996, when the complete genome of the first photosynthetic organism, the cyanobacterium *Synechocystis sp.* PCC6803 [Kaneko, et al., 1996] was sequenced. In the following years, several cyanobacteria and other, even non-photosynthetic bacteria revealed putative phytochrome homologous genes as part of their genomes.

#### 2.1.4 Phytochrome signaling

For a long period of time, the generally accepted dogma on the function of phytochromes was that they are soluble, cytosolic proteins that interact with as yet unknown components to forward their signal to nuclear genes [Quail, 1991]. This picture has changed dramatically, when in 1996 and 1999, results were published that demonstrated unequivocally that photoactivated phy molecules themselves, are translocated from the cytosol to the nucleus [Kircher, et al., 1999, Kleiner, et al., 1999, Sakamoto and Nagatani, 1996, Yamaguchi, et al., 1999]. In the following year, this was shown for all five phytochromes from higher plants, supporting the generality of the mechanism [Nagy and Schäfer, 2000a, Nagy and Schäfer, 2000b] which then became widely accepted. The search for a putative second messenger interacting with phys was superseded with the aim to identify molecules that could, together with phys, play a part in nuclear transcriptional regulatory complexes. The first interacting factor to be identified in 1998 by means of two-hybrid screens, was the phytochrome interacting factor 3 [PIF3]. This molecule is a member of the well known basic helix-loop-helix superfamily of transcriptional regulator proteins [Ni, et al., 1998] and binds to both, PhyA and PhyB after activation by light, but with different affinities [Ni, et al., 1999, Zhu, et al., 2000]. Subsequently, a G-box DNA sequence GACGTG, could be identified as the motif to which PIF3 specifically binds. This G-box motif was found to be present in several light regulated promoters such as those regulating the expression of CIRCADIAN CLOCK-ASSOCIATED PROTEIN1 [CCA1] and LATE ELONGATED HYPOCOTHYL [LHY] [Martinez-Garcia, et al., 2000]. The latter two themselves, encode transcription factor related proteins known to be involved in the regulation of the expression of CHLOROPHYLL A/B BINDING PROTEIN [CAB] gene and the circadian clock [Green and Tobin, 1999, Schaffer, et al., 1998, Wang, et al., 1997, Wang and Tobin, 1998]. It was thus proposed that PIF3 might act as a central control point through which phytochromes control both, photomorphogenesis as well as the circadian clock. And as results, published by Ni et al. and Martinez-Garcia et al. [Martinez-Garcia, et al., 2000, Ni, et al., 1999] showed that light activated PhyB binds specifically to DNA-bound PIF3, it was proposed that phytochromes might act as light activated components of transcription regulator complexes that are synthesized and reside as the inactive form in the cytosol and become translocated to the nucleus upon photoconversion to the active form [Martinez-Garcia, et al., 2000, Quail, 2000]. Figure 2.4 illustrates the

general mechanism by which phytochromes interact with PIF3 and activate transcription.

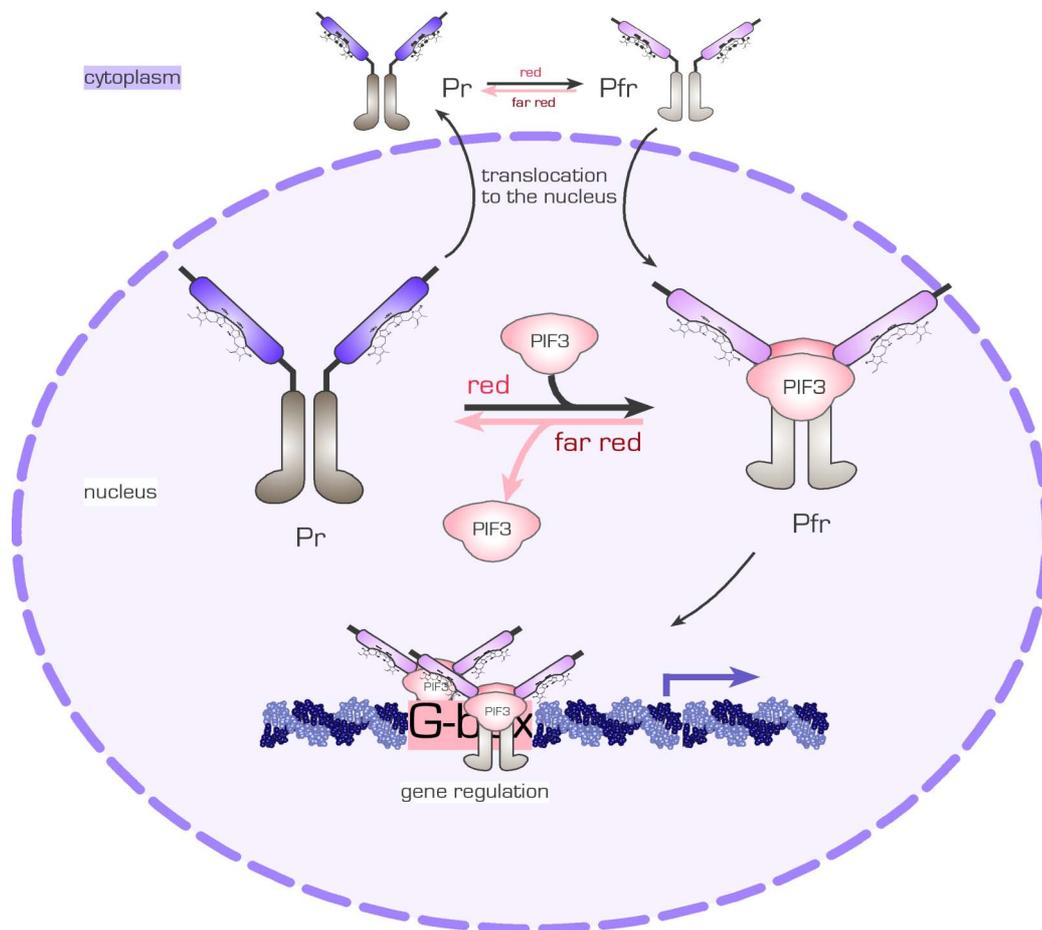


Figure 2.4. **Plant phytochrome signaling model.** A simplified mechanism by which phytochromes are believed to directly alter gene expression is illustrated. As postulated, direct targeting of light signals to promoter regions is accomplished by the light dependent interaction of phytochromes ( $P_r$  and  $P_{fr}$ ) with the basic helix-loop-helix factor PIF3. The inactive  $P_r$  form of phy is activated by a light stimulus ( $P_{fr}$  formation) and translocates from the cytoplasm in to the nucleus. In the  $P_{fr}$  form, phy binds to free as well as DNA bound PIF3. The latter binds to G-box promoter regions in a sequence specific manner. The G-box (CACGTG, red) is present in various light-regulated promoters and the phy:PIF3 complex is proposed to be part of a pre-initiation complex for transcription.

In the time following, several other phytochrome interacting proteins were identified, e.g. NDPK2, PSK1 [Quail, 2000] and ARR4 [Fankhauser, 2002, Sweere, et al., 2001], that specifically interact with either PhyA or PhyB or with both, indicating a possible common signaling pathway for different phytochromes. Additionally, a long suggested model has been proven *in vitro*, namely the ability of phys to autocatalytically phosphorylate an intramolecular serine residue and thus, the fact that phys constitute light regulated

protein kinases of the eukaryotic Ser/Thr/Tyr class [Fankhauser, et al., 1999, McMichael and Lagarias, 1990, Wong, et al., 1986, Yeh and Lagarias, 1998]. Although recombinant PhyA was shown to phosphorylate PHYTOCHROME-KINASE SUBSTRATE 1 (PKS1) in a light dependent manner [Fankhauser, et al., 1999], the way how the phy kinase activity is involved in signal transduction pathways remains controversial.

The signaling mechanism in the recently discovered bacterial phytochrome systems is by far more well-defined than in plants. From sequence analysis it soon became obvious that bacterial phytochromes are composed of two structural domains, an N-terminal sensor domain with high sequence similarity to plant phytochromes, and a C-terminal signaling domain that carries the conserved motifs of sensor histidine kinases of the bacterial two-component systems. This system is well known from bacterial chemotaxis in *E. coli* and is basically composed of as few as two components, a receptor histidine kinase and a dedicated response regulator [Falke, et al., 1997, Stock, et al., 2000]. As illustrated in Fig. 2.5, many bacterial phytochromes are organized together with their cognate response regulator in the same operon which facilitates the identification of related components.

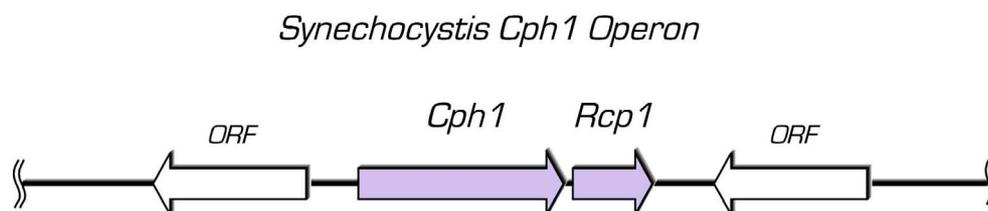


Figure 2.5. **Genomic organization of the *Cph1* operon.** The *Cph1* operon in *Synechocystis sp.* PCC6803 is a two-member transcriptional unit encoding the phytochrome-like *cph1* and its cognate response regulator *rcp1*. A similar organization can be found in several other organisms e.g. in *Calothrix sp.* PCC7601 (not shown).

The general mechanism by means of which these bacterial photoreceptors generate a cellular signal was demonstrated *in vitro* for the cyanobacterial phytochromes *cph1* from *Synechocystis sp.* PCC6803 [Yeh, et al., 1997] and for *cphA* and *cphB* from *Calothrix sp.* PCC7601 and their interacting response regulators *RcpA* and *RcpB* [Hübschmann, et al., 2001]. The authors showed that bacterial phytochromes indeed auto-phosphorylate a conserved histidine in a light dependent manner, and that this phosphate group is subsequently transferred to a conserved aspartate side chain in the response regulator, just as claimed by the two-component principle. Figure 2.6 shows a simplified model for the light regulated action of bacterial phytochromes from *Calothrix sp.*

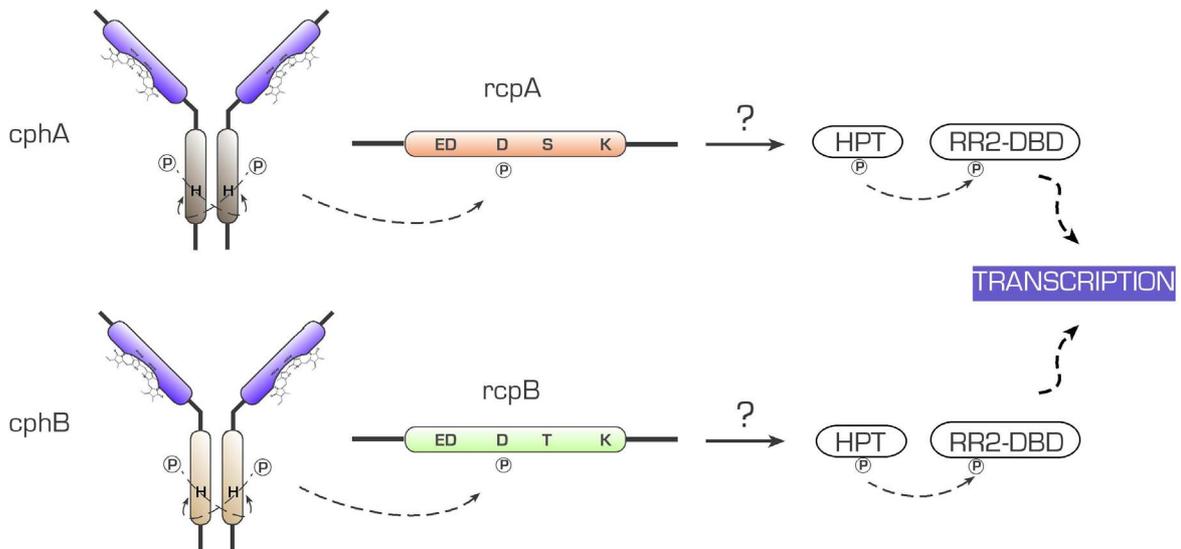


Figure 2.6. **Bacteriophytochromes**, here cphA and cphB from the cyanobacterium *Calothrix sp.* PCC7601, are dimeric receptor histidine kinases with an open chain tetrapyrrole chromophore bound to a conserved region in the sensor domain (blue). The absorption of light triggers a conformational change that is transferred to the histidine kinase domain in an unknown fashion. The thus activated kinase cross-autophosphorylates a histidine side chain (H) in the dimer and then transfers the phosphoryl group to an aspartate (D) of a cognate response regulator molecule (RcpA or RcpB in *Calothrix sp.*). The activated response regulator transfers the phosphate to a conserved histidine in a histidine phosphotransferase (HPT) which donates it to a second response regulator (RR2) fused to an output domain (e.g. DNA binding domain DBD). At present, neither a HPT nor a RR2 have been identified for any bacterial phytochrome.

In *Calothrix sp.*, the response regulators of phytochromes are proteins built of approx. 140 amino acids and contain the essential residues conserved in the CheY superfamily of response regulators: a triad of acidic residues (E17, D18, D70 in RcpA) a serine or threonine (S100) and a lysine (K122). Even though they possess a C-terminal extension of seven or eight amino acids and an extended loop connecting helix 2 and beta strand 3 with respect to CheY, they contain no functional output domain. It was thus proposed by Vierstra and Davis (Vierstra and Davis, 2000) that after receiving a phosphate from the activated receptor molecule, the response regulator transfers this group to additional downstream partners in order to forward the signal. These could involve a histidine phosphotransferase (HPT) which accepts an aspartate-bound phosphate group on a histidine side chain and subsequently phosphorylates a second response regulator that is fused to an output domain, e. g. a DNA-binding domain to serve as a transcription factor and affect gene expression, or a protein-protein interaction site. Although in the case of *Calothrix* the two phy associated response regulators do not exhibit a distinct output domain as judged from sequence and secondary structure predictions, two possible alternatives for the signal forwarding are most probable. Either, the phospho-

relay continues to a HPT followed by a second response regulator with an output domain and constitutes a four step phosphorelay with His-Asp-His-Asp phospho-transfer, or the RR itself possesses the ability to interact directly with an as yet unidentified target protein (like in the case of CheY that directly interacts with the motor switch component FliM, see 5.1.1) and allosterically alters cellular functions like motility.

## 2.2 Conceptual Formulation

This work focuses on the general principle underlying the phytochrome mediated light perception in plants and bacteria from a structural point of view. Phytochromes have been under intense investigation since their discovery more than forty years ago. Nevertheless, the atomic structures of the receptors or interacting factors have not yet been revealed and a profound understanding of the mechanism of light absorption and signal transduction remains a great challenge. Due to the recent discovery of a new group of bacterial phytochromes in cyanobacteria and other microorganisms, this field has experienced an immense impulse offering new approaches towards a comprehensive insight into light signaling.

The central aim of this study was to elucidate the atomic structure of components of the phytochrome signaling network by X-ray crystallographic methods. Among the selected targets were two cyanobacterial response regulators, RcpA and RcpB, from the cyanobacterium *Calothrix sp.* These proteins specifically interact with their cognate bacterial phytochromes CphA and CphB, and, being homologous to receiver modules of the CheY type, relate the bacterial light sensing to the common bacterial two-component signaling systems. The availability of structural information should permit a detailed comparison with other two-component response regulators and explain the function of differences in the primary structure of phytochrome associated response regulators and other characterized homologues. Besides this, two recombinant fragments of the oat phytochrome PhyA were selected for characterization and X-ray structural analysis. The two fragments, a 59 and a 65 kD N-terminal domain of the full length receptor (see fig. 2.7) comprise the chromophore bearing part that is responsible for the initial interaction with light. Knowledge of the structure down to the atomic level would clarify the interaction of the bilin chromophore with its proteinaceous environment and, for the first time give structural insight into the principles underlying the most prominent light receptor, besides the visual pigments of vertebrates.

The two receptor fragments were selected for several reasons. Firstly, they comprise fully photoconvertible entities with the spectral characteristics comparable to those of full length phytochromes. Therefore, they can be regarded as fully functional receptor domains and information obtained for them should equally apply for the unmodified receptors. The fragments were obtained by tryptic digest of full length oat PhyA and are believed to be single domain molecules that would have the obvious advantages over multidomain proteins, like e.g. a more restrained conformational flexibility. These constructs were cloned and heterologously expressed in yeast expression systems in previous works (Gärtner, et al., 1996, Mozley, et al., 1997), and high amounts of

functional protein were accessible, which is necessary for crystallization trials. For the latter purpose, the two proteins were to be expressed in the yeasts *Pichia pastoris* (59 kD PhyA) and *Hansenula polymorpha* (65 kD PhyA) by fermentation and a purification protocol was to be established that allows the gentle enrichment of recombinant proteins from cell extracts with the special requirements posed by crystallization (high purity, homogeneity, functional integrity, conformational integrity and others). The receptor domains were to be characterized spectroscopically and physicochemically and their suitability for crystallization was to be addressed. Furthermore, the quality of the proteins and the dynamics involved in the photoconversion should be investigated by NMR techniques, including one-dimensional  $^1\text{H}$  NMR and two-dimensional  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear magnetic resonance.

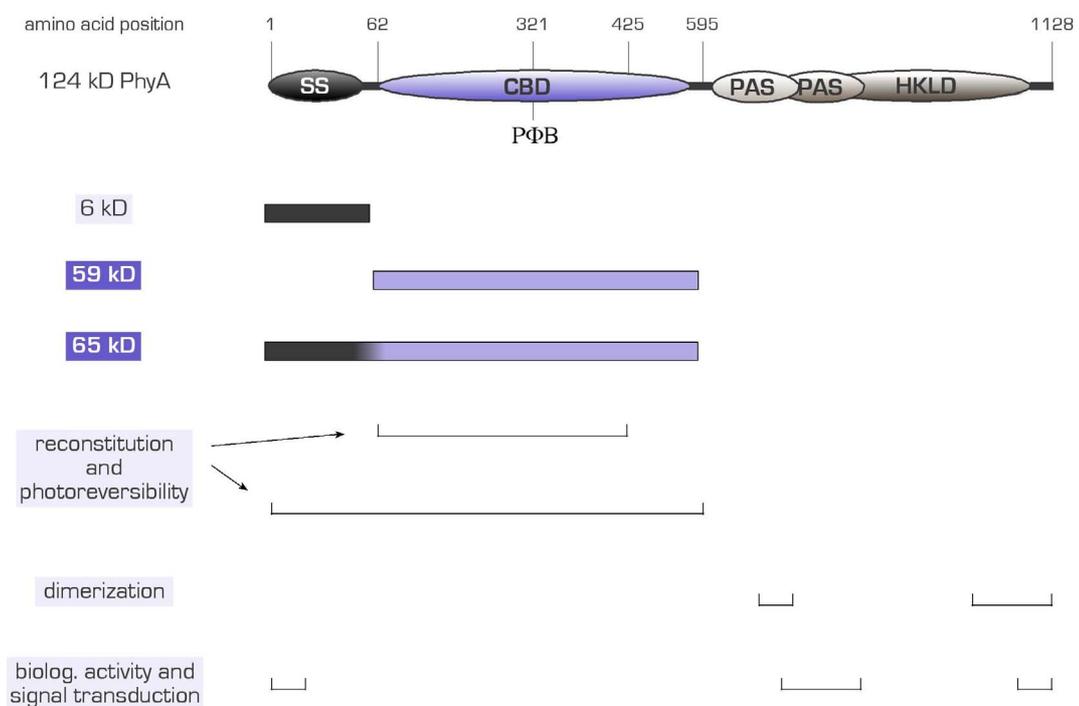


Figure 2.7. **Functional entities in *A. sativa* PhyA.** Shown is a schematic representation of the organization of main functional domains in oat PhyA. The N-terminal half of the protein contains a serine-rich stretch (SS) and the sensor domain (CBD) that binds the phytochromobilin chromophore (PΦB). The C-terminal half harbors two Per-Arnt-Sim domains (PAS) and a histidine kinase-like domain (HKLD). Two constructs, 59 and 65 kD PhyA, are indicated (blue). They represent fully photoconvertible receptor domains and were selected as targets for physicochemical characterization and crystallization.

## 3 Materials and Methods

### 3.1 Materials

All chemicals and reagents used throughout this work were of reasonable purity and purchased from either Sigma (Deisenhofen), Merck (Darmstadt), ICN (Eschwege) or Hampton Research (Laguna Niguel, USA) unless otherwise indicated. All solutions and lab ware for bacterial/yeast cell cultures and molecular biological techniques were autoclaved or sterile filtered. All chemicals used for protein crystallization trials were of the highest purity available, and all solutions were sterile filtered before use. Information on specific chemicals, reagents or lab ware is given below in the respective instruction and in the list appended to this paragraph, if reasonable. All experiments where water was used, the term water refers to deionized tap water from the Institute's deionized water supply. All numbers given in *per cent* [%] refer to either weight per volume or, in cases of liquids, to volume per volume. All experiments involving standard techniques were performed following the manufacturers protocols and recommendations, when available, if not otherwise indicated.

#### 3.1.1 Equipment

##### Instruments

Chromatography media	Dimethyl ethyl cellulose DE52, Whatman, Fairfield, NJ, USA; Ni-NTA, Qiagen, Hilden; Hydroxyapatite BioRad, Richmond, CA, USA; all other media were purchased from Amersham Pharmacia Biotech, Freiburg
CD spectrometer	JASCO J-715 Spectropolarimeter, Easton, MD, USA
DNA sequencer	ABI Prism 377, Applied Biosystems, Foster City, CA, USA
FPLC System and Accessories	Amersham Pharmacia Biotech, Freiburg
French Press	Aminco 20K French Pressure Cell, Polytec GmbH, Waldbronn
Incubator	Multitron AJ120, Infors AG, Bottmingen, Switzerland
PCR machine	GeneAmp PCR System 9700, Applied Biosystems, Foster City, CA, USA
Phosphoimager	FujiFilm FLA-2000, Fuji Photo Film Co., Ltd., Tokyo, Japan
Sonifier	Sonifier Cell Disruptor B-30, Branson Sonic Power Co. Danbury, CT, USA
Transfection apparatus	Gene Pulser, BioRad, Richmond, CA, USA
X-ray source	Rigaku Rotating Anode Generator, Rigaku, The Woodlands, TX, USA

**Consumables**

Centriprep	Centriprep K, Millipore, Eschbronn
Crystallization consumables	Hampton Research, Laguna Niguel, CA, USA
Cuvette for CD spectroscopy	Hellma GmbH & Co KG, Müllheim
Dialysis Tubing	10.000 Da exclusion limit, Sigma-Aldrich Chemie, Deisenhofen
Electroporation cuvettes	0.2 mm gap, BioRad, Richmond, CA, USA
Filter paper	Whatman 3MM, Whatman, Fairfield, NJ, USA
Microfuge Tubes	0.5 and 1.5 ml, Eppendorf AG, Hamburg
PVDF Membrane	Immobilon P, Millipore
Spin columns	AutoSeq G-50, Amersham Pharmacia Biotech, Freiburg and MultiScreen-HV 96, Millipore, Eschbronn
X-ray film	XAR-5, Eastman Kodak GmbH, Stuttgart-Wangen

**Kits**

DNA purification kit	QIAquick, Qiagen, Hilden
DNA sequencing kit	ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit, Perking Elmer, Wellesley, MA, USA
Plasmid isolation kit	QIAprep, Qiagen, Hilden
Quick change site directed mutagenesis kit	Stratagene, La Jolla, CA, USA
RNEasy RNA kit	QIAquick, Qiagen, Hilden

**Enzymes**

DNA Ligase	T4 DNA Ligase, Gibco BRL, Invitrogen GmbH, Karlsruhe
DNaseI	Roche Diagnostics GmbH, Mannheim
DNA polymerase	TaKaRa LA Taq, Takara shuzo Co., Ltd., Shiga, Japan Herkulase enhanced DNA polymerase, Stratagene, La Jolla, CA, USA
<i>Pfu</i> DNA Ligase	Stratagene, La Jolla, CA, USA
Polynucleotide kinase	USB, Cleveland, OH, USA
Restriction endonucleases	New England Biolabs GmbH, Schwalbach/ Taunus
Shrimp alkaline phosphatase	USB, Cleveland, OH, USA
Thrombine	Amersham Pharmacia Biotech, Freiburg

**Fine Chemicals**

Agarose	Seakem, Teknova, Half Moon Bay, CA, USA
Coomassie Brilliant Blue	Serva, Heidelberg
DNA size standard	One kb DNA size marker, Gibco BRL, Invitrogen GmbH, Karlsruhe
dNTPs	PCR nucleotide mix, Amersham Pharmacia Biotech, Freiburg
Gel filtration size standard	Gel filtration LMW and HMW Calibration kits, Amersham Pharmacia Biotech, Freiburg
IPTG	Gerbu Biotechnik GmbH, Gaiberg
L-Selenomethionine	Sigma-Aldrich Chemie, Deisenhofen
PMSF	Sigma-Aldrich Chemie, Deisenhofen
Protease inhibitor cocktail	Complete Protease Inhibitor Cocktail, EDTA free, Hoffmann-La Roche Ltd., Basel, Switzerland
Protein MW marker	LMW, Amersham Pharmacia Biotech, Freiburg Kaleidoscope Prestained Standard, 161-0325 and 161-0324, BioRad, Richmond, CA, USA Prestained Protein Marker, broad range (6.5, 17, 25, 33, 48, 62, 83, 175 kD; New England Biolabs
SDS, 99%	Roth, Karlsruhe
Western Blocking Reagent	Hoffmann-La Roche Ltd., Basel, Switzerland
Zeocin	Invitrogen GmbH, Karlsruhe

**Antibodies**

Primary antibody	Mouse anti His <sub>6</sub> IgG, Dianova, Hamburg
Secondary antibodies	Alkaline phosphatase conjugated goat anti mouse IgG, Jackson Immuno Research Laboratories Inc., West Grove, PA, USA, purchased from Dianova, Hamburg

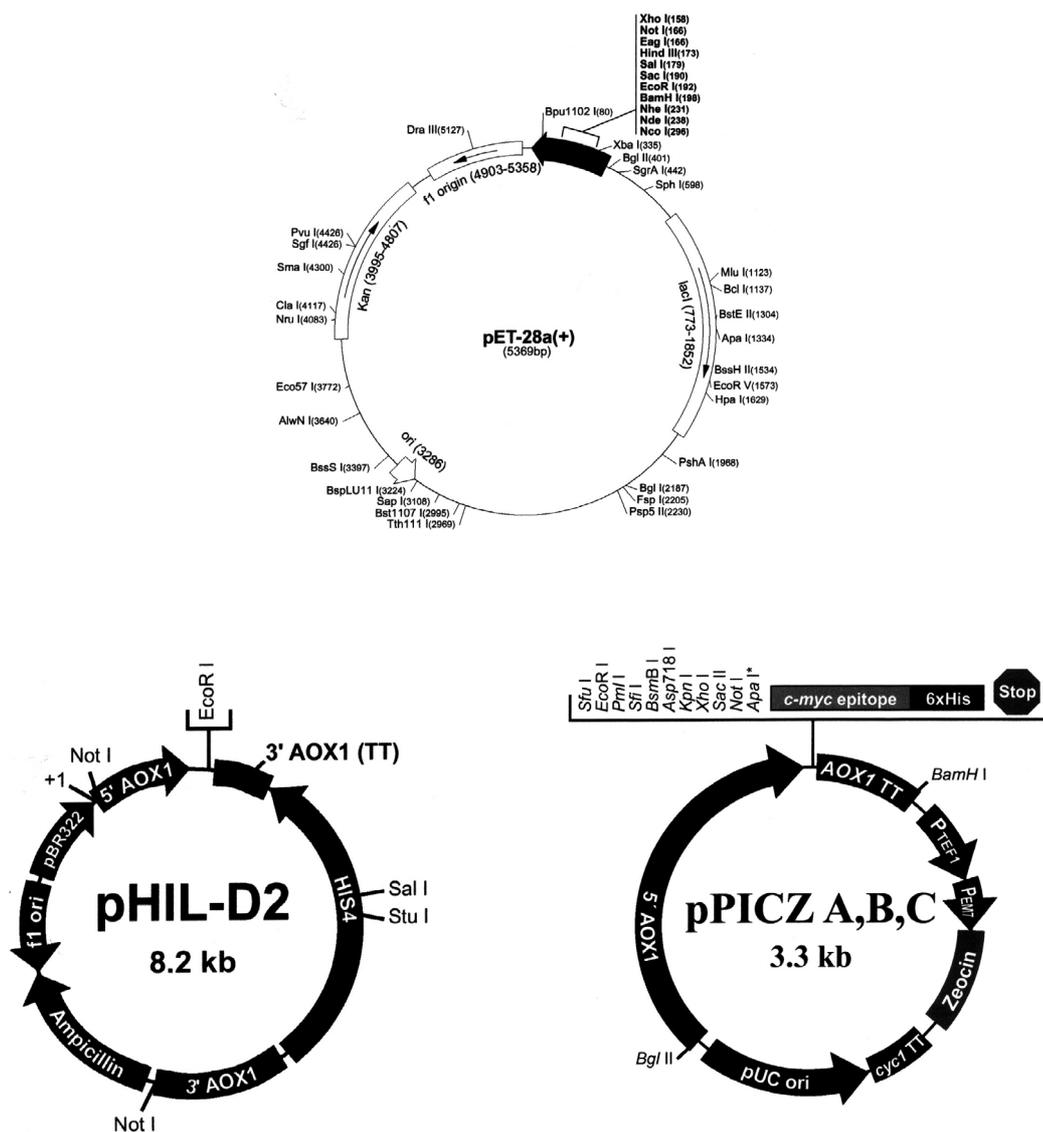
**3.1.2 Bacterial and yeast strains**

Unless indicated otherwise, *Escherichia coli* strain DH5 $\alpha$  was used for plasmid growth and maintenance, and strain BL21(DE3) [Studier, et al., 1990] was used for high-level expression of recombinant proteins directed by plasmid pET28a(+) [Novagen] derived vectors harboring the gene of interest under the control of the T7 promoter. For the expression of L-selenomethionine labeled proteins the methionine auxotroph *Escherichia coli* strain B834(DE3) [Novagen] was used.

For eukaryotic protein expression the yeasts *Hansenula polymorpha* (Rhein Biotech, Düsseldorf) and *Pichia pastoris* GS115 *his4* (Invitrogen) were used.

### 3.1.3 Plasmids

For bacterial protein expression, all plasmids were constructed by ligating appropriately digested PCR products into plasmid pET28a(+) [Novagen] which placed a His<sub>6</sub>-tag followed by a thrombin protease recognition site at the amino terminus of the protein. For heterologous expression in yeast the plasmids pHILD2 and pPICZ A [Invitrogen] were combined with PCR derived DNA fragments encoding the respective protein and a C-terminal thrombin recognition site followed by a His<sub>6</sub>-tag. All resulting plasmids were sequenced to ensure no errors were introduced into the gene during the amplification reactions.



## 3.2 General Molecular Biological Techniques

### 3.2.1 Culture of *E. coli* DH5 $\alpha$ cells for plasmid growth

In a conical flask, 50 ml of LB broth (10 g Tryptone, 5 g NaCl, 5 g Yeast Extract per liter; [Sambrook, et al., 1989]) with an appropriate antibiotic were inoculated with a single colony of *E. coli* DH5 $\alpha$  harboring the plasmid of interest. The culture was incubated overnight on a platform shaker at 37°C and 250 rpm. From this culture the cells were harvested by centrifugation (3000xg<sub>max</sub>, 5 min) and the plasmid DNA was purified as described.

### 3.2.2 Purification of plasmid DNA from *E. coli* cells

All plasmids used throughout this work were purified from an overnight *E. coli* DH5 $\alpha$  cell culture using the Qiaprep® (Qiagen) Miniprep protocol as recommended by the manufacturer. The bound plasmid DNA was eluted with 50  $\mu$ l elution buffer per Miniprep.

### 3.2.3 Isolation of total RNA from *Arabidopsis thaliana*

For total RNA isolation from plant tissue the RNeasy® Plant Mini Kit (Quiagen) was used. *Arabidopsis thaliana* *l. er.* plants were grown from seeds in plastic ware on the laboratory bench. Sterile seeds were spread on autoclaved soil and grown for approx. 2 weeks with additional supply of artificial white light issued from two neon lamps (Sylvania GRO-LUX, F15W/GRO for plant growth) at an average distance of 75 cm, and occasional watering. A maximum of 100 mg of the seedling's green parts were harvested and ground in a mortar under liquid nitrogen. The nitrogen was allowed to evaporate and the remaining powder was immediately suspended in 450  $\mu$ l RLT buffer by vigorously vortexing. All subsequent steps were carried out according to the manufacturers protocol. The final elution step was performed with 30  $\mu$ l of RNase free water. The RNA solutions were stored at -20°C until further use.

### 3.2.4 General PCR protocol for DNA amplification

All DNA amplification reactions were performed as follows: in a PCR reaction cap, an appropriate amount of template DNA (generally 100 nl of a standard Miniprep (Qiagen) DNA preparation) was added to a mixture of 5  $\mu$ l of each oligonucleotide primer (10  $\mu$ M), 10  $\mu$ l 10 mM dNTPs (2.5 mM each), 5  $\mu$ l 10x Herculase reaction buffer, one  $\mu$ l DMSO, 1  $\mu$ l Herculase Polymerase (Invitrogen) and ddH<sub>2</sub>O to a final volume of 50  $\mu$ l. After an initial denaturation step at 94°C for 3 min, the reactions were cycled (denaturation, annealing and extension) for 30 cycles at 94°C for 45 sec, 45°C for 45 sec and 63°C

for 2-3 min. The reaction process was monitored by applying aliquots of the reaction mixture on an agarose gel.

### 3.2.5 Reverse transcription polymerase chain reaction (RT-PCR)

In order to generate cDNA from RNA, the SIGMA Enhanced Avian RT-PCR Kit (Sigma) was used in the following reaction setup: 5  $\mu$ l aqueous total RNA solution, prepared as described (3.2.3), were combined with 1  $\mu$ l Desoxyribonucleotide-Mix (10 mM each), 2.5  $\mu$ l primer mix (forward and reverse primer, 2.5  $\mu$ M each, see below) and 8.5  $\mu$ l water in a standard PCR tube. After incubation for 10 minutes at 75°C, the reaction mix was cooled on ice and supplemented with 2  $\mu$ l 10x AMV-RT buffer, 1  $\mu$ l AMV reverse transcriptase and 0.5  $\mu$ l RNase inhibitor (supplied with the kit). After incubation for 50 minutes at 42°C, the mixture was again cooled on ice. For the amplification of cDNA in a PCR reaction, 5  $\mu$ l from the RT reaction mix served as template in a standard PCR as described (3.2.4).

### 3.2.6 Purification of DNA after enzymatic reactions

For the purification of DNA after enzymatic reactions the Qiaquick® (Qiagen) procedure was used essentially as recommended by the manufacturer, except that the column was washed twice with washing buffer and the purified DNA was eluted from the columns with 120  $\mu$ l of elution buffer in all cases where the DNA was further processed in enzymatic reactions. In all other cases the DNA was eluted with 30  $\mu$ l elution buffer to ensure a high DNA concentration.

### 3.2.7 Preparation of vector DNA for ligation reactions

Vector DNA was prepared as described under 3.2.1 and 3.2.2. For restriction of vector DNA, 120  $\mu$ l Miniprep™ DNA solution was digested with 100 units of each of the appropriate restriction endonucleases in a total volume of 150  $\mu$ l supplemented with reaction buffer as recommended by the manufacturer. After 30 min at 37°C, the DNA was purified from the reaction mixture as described under 3.2.6. The 5'-phosphate group was removed from the linearized DNA molecule by shrimp alkaline phosphatase treatment as follows: 15  $\mu$ l reaction buffer and 3 units shrimp alkaline phosphatase were added to the DNA solution in a final volume of 140  $\mu$ l. Incubation was for 30 min at 37°C followed by purification of the vector DNA from the reaction mixture as described (3.2.6) and stored at -20°C until further use.

### 3.2.8 Preparation of PCR-derived DNA fragments for ligation reactions

DNA fragments to be ligated into expression vectors were digested with the appropriate restriction endonucleases to produce cohesive ends as follows: 120  $\mu$ l DNA preparation (3.2.4), 20  $\mu$ l of the restriction endonuclease reaction buffer and 100 units of each restriction endonuclease were mixed in a total volume of 150  $\mu$ l. After incubation at 37°C for 30 min, the DNA was purified from the reaction mixture as described under 3.2.6.

### 3.2.9 Ligation of DNA molecules

Linearized vector DNA and PCR-derived DNA fragments were prepared as described under 3.2.7 and 3.2.8, respectively. The ratio between linearized vector and insert DNA in a reaction was ideally 3:1, roughly estimated through agarose gel analysis. All ligation reactions were prepared by mixing 4  $\mu$ l 5x ligation buffer (Gibco BRL), 2  $\mu$ l insert DNA, 1  $\mu$ l vector DNA, 1  $\mu$ l 10 mM ATP and 2  $\mu$ l T4 DNA ligase (Gibco BRL) to 10  $\mu$ l water in a standard PCR tube. The reaction mixtures were inserted in a PCR thermocycler and, after an initial step of 24°C for 15 min, cycled for 99 cycles of 24°C for 2 min, 18°C for 2 min and 12°C for 2 min followed by 1 h at 12°C. This reaction mixture was used for transforming bacterial cells (3.2.18) without further modification.

### 3.2.10 Restriction analysis of plasmid DNA

Plasmid DNA isolated from transformed bacterial cells (3.2.1 and 3.2.2), was generally analyzed via agarose gel electrophoresis after restriction digestion. The plasmid DNA was digested with the restriction endonuclease(s) used to construct the desired plasmid by mixing 5  $\mu$ l plasmid preparation, 10 units of each restriction endonuclease in a total volume of 10  $\mu$ l of 1x reaction buffer supplied with the restriction endonuclease. After incubation at 37°C for 30 min the DNA was analyzed on an agarose gel as described.

### 3.2.11 Analysis of DNA by agarose gel electrophoresis

Agarose gel electrophoresis of DNA was performed in commercially available submarine gel tanks of appropriate size following the method of (McDonnell, et al., 1977) and (Southern, 1979). All agarose gels used throughout this work were run in TAE buffer (50 mM Tris acetate, 1 mM EDTA, pH 8.0) and were prepared by dissolving 1.0 % agarose and 0.5  $\mu$ g ml<sup>-1</sup> ethidium bromide in TAE buffer. Samples containing an appropriate amount of DNA were mixed with 3x loading buffer (0.1% bromophenol blue, 15% Ficoll type 400 in water) prior to sample application. The gels were run at 4 V cm<sup>-1</sup> until optimal separation was achieved. The DNA was visualized via fluorescence

excitation by illumination with UV light (302 nm). A 1 kb ladder (Gibco BRL) was applied as a size standard.

### 3.2.12 Sequencing of plasmid DNA

All sequencing reactions were carried out following the chain termination method of Sanger [Sanger, et al., 1980, Sanger, et al., 1977], using ABI Prism BigDye™ Terminator Cycle Sequencing Ready Reaction Kit (Perkin Elmer) as recommended by the manufacturer. In brief, 2.5 µl plasmid DNA, 7.5 µl water, 1 µl DMSO, 5 µl sequencing primer (10 µM) and 4 µl BigDye™ reagent were mixed in a standard PCR tube. The reaction was cycled [denaturation, annealing and extension] after an initial 3 min denaturation step at 94°C for 30 cycles [45 sec at 94°C, 45 sec at 43°C and 5 min at 60°C]. Then, 10 µl water were added to the reaction mixture, and the DNA was purified from other components of the reaction mixture by size exclusion chromatography on Sephadex G-50 columns as follows: the DNA solution was applied to either an AutoSeq G-50 disposable spin column (Amersham Pharmacia Biotech) or to a Millipore MultiScreen-HV 96 well plate. Subsequently, the DNA was eluted from the columns as recommended by the respective manufacturer, dried in a speed vac and resuspended in 3 µl denaturation buffer [100 mg dextran blue and 1 ml 25 mM EDTA in 7 ml deionized water]. The DNA probes were denatured for 2 min at 96°C and the amplification products were separated and analyzed in an ABI Prism 377 DNA sequencer.

### 3.2.13 Introduction of point mutations by site-directed mutagenesis

To introduce point mutations into genes cloned in expression vectors, the QuickChange™ site-directed mutagenesis kit (Stratagene) was used essentially as recommended by the manufacturer. Primers were designed such that the respective mutations were flanked by approx. 30 complementary bases on either side to guarantee specific annealing. Successful mutagenesis was approved by DNA sequence analysis [3.2.12]

### 3.2.14 Construction of mutant genes by PCR

In cases where the DNA sequence to be altered was a gene's 5´- or 3´- end [e.g. introduction of a His<sub>6</sub>-tag, a thrombin cleavage site or new restriction sites] the novel gene sequence was introduced into the gene by PCR with an overlapping oligonucleotide primer bearing the mutation. The oligonucleotide was designed to cover approx. 30 complementary base pairs [bp] down- or upstream from the desired mutation, the sequence to be introduced and flanking sequences recognized by restriction

endonucleases for the generation of overlapping ends. Subsequently, the resulting PCR fragment was cloned into an expression vector as described.

### 3.2.15 Isolation of DNA fragments from agarose gels

After separation by agarose gel electrophoresis, DNA fragments were cut out from the gel and extracted by using the QIAquick® Gel Extraction Kit™ (Qiagen, Germany) according to the supplier's protocol.

### 3.2.16 Determination of DNA concentration in aqueous solution

The concentration of DNA in aqueous solutions was determined by measuring the UV absorption at 260 and 280 nm in quartz cuvettes with 0.5 cm path length. For pure double stranded DNA the absorption ratio  $A_{260}/A_{280}$  is approx. 2. Concentrations were calculated as follows:

double stranded DNA [ $\mu\text{g ml}^{-1}$ ] =  $44 \times A_{260} \times \text{dilution factor}$ .

### 3.2.17 Preparation of electrocompetent *E. coli* cells

A small amount of the frozen cell suspension from a glycerol culture of the requested *E. coli* strain, stored at  $-80^{\circ}\text{C}$  was removed with a sterile platinum loop. The cells were stroke out on a SOC [Sambrook, et al., 1989] agar plate and incubated overnight at  $37^{\circ}\text{C}$ . At the following day, 50 ml LB broth were inoculated with a single bacterial colony and the flask was incubated overnight on a platform shaker at  $37^{\circ}\text{C}$ , 250 rpm. This starter cell culture was thereafter diluted 100-fold in LB broth, and the cells were further incubated on a platform shaker at 250 rpm,  $37^{\circ}\text{C}$ . Typically, one liter LB broth in a two liter Erlenmeyer flask was used to grow the bacteria, and all following quantities refer to a one liter cell culture. When the bacterial culture reached an  $\text{OD}_{600}$  of approx. 0.8, the flask was removed from the shaker and cooled down to  $0^{\circ}\text{C}$  in an iced water bath. After 20 min, the bacteria were harvested by centrifugation in 450 ml centrifuge tubes at  $4^{\circ}\text{C}$  for 10 min at  $2500 \times g_{\text{max}}$ . The cell pellet was washed with 450 ml ice cold water and the cells were recovered by centrifugation as described above. This washing step was repeated twice with first water followed by an ice cold solution of 10% glycerol. Finally, the cells were resuspended in 2.5 ml of ice cold 10% glycerol, dispensed as 50  $\mu\text{l}$  aliquots into microcentrifuge tubes, shock frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until further use.

### 3.2.18 Transformation of *E. coli* cells

For the transfer of DNA into electrocompetent *E. coli* cells as prepared in 3.2.17, the cells were thawed on ice. In all cases where DNA from plasmid preparations (Miniprep,

Qiagen] was to be transferred, 75  $\mu$ l of electrocompetent *E. coli* cells were mixed with 100 nl of plasmid DNA solution obtained as described under 3.2.2. If DNA from ligation reactions was to be transferred, 10  $\mu$ l of the ligation reaction mix was combined with 75  $\mu$ l electrocompetent *E. coli* cells. The cells were transferred to ice cold electroporation cuvettes (0.2 mm gap width, Bio-Rad) and incubated on ice for 5 minutes. Electroporation was performed using a Biorad Gene Pulser System (Bio-Rad) with the settings for *E. coli*: charging voltage 1.5 kV, resistance 800  $\Omega$ , capacitance 25  $\mu$ F. The transformed cells were immediately resuspended in 1 ml LB medium, transferred to 10 ml plastic cell culture tubes, and incubated for phenotypal gene expression in a shaker for 45 minutes at 37°C. Bacteria were plated out (50  $\mu$ l in all cases where plasmid DNA preparations were to be transformed, and 300  $\mu$ l in all other cases) on LB agar plates containing the appropriate antibiotic and incubated overnight at 37°C, plates were always kept upside down. If plasmid DNA for *Pichia p.* (pPICZ A) was transformed, low-salt LB medium (10 g Tryptone, 5 g NaCl, 5 g Yeast Extract) and low-salt LB agar plates, both containing 25  $\mu$ g ml<sup>-1</sup> Zeocin™ were used.

### 3.2.19 Cloning of *PIF3* from *A. thaliana*

The *PIF3* gene [accession number AF100166] from *A. thaliana* *l.e.* was amplified by standard PCR (3.2.4) from cDNA obtained as described in 3.2.5, using the following oligonucleotide primers, designed to introduce the *Xho* I (fw) and *Nde* I (rev) restriction sites:

PIF3NdeI fw

5'-ATC TTA CAT ATG ATG CCT CTG TTT GAG CTT TTC AGG CTC ACC -3'

PIF3XhoI rev

5'-TCT TCT GAT CAG TTT TGT GGA TCG TCG TGA CTC GAG ATC TTA-3'

The PCR product was processed (3.2.6) and ligated into the vector pET28a(+) for expression in *E. coli*, as explained in 3.2.7 through 3.2.11. In order to verify the correctness of the vector product, the following oligonucleotide primer were used for DNA sequencing (3.2.12):

PIF3NdeI fw

PIF3XhoI rev

SeqPIF3 500 5'-TAG AGC TAG AGA TCT TCC TAG -3'

seqPIF3 1000 5'-GAA CGG GTT TGG GTT CAA AG -3'

In order to clone the *PIF3* gene into the yeast vector pPICZ A for expression in yeast, the gene was amplified by PCR from an *A. thaliana* cDNA library (courtesy of W. Frommer,

ZMBP, Tübingen) and processed as described. The oligonucleotide primers used were designed to introduce an *EcoR* I (fw) and a *Not* I (rev) restriction sites, a 5' yeast consensus sequence (Romanos, et al., 1992) for initiation of transcription (ACA ATA **ATG** CCT) and a sequence encoding a C-terminal thrombin cleavage site (L V P R G S) upstream to the His<sub>6</sub>-tag supplied by the vector.

PIF3pPICZfw EcoRI

5'- AAA CGA GAA TTC ACA ATA ATG CCT CTG TTT GAG CTT TTC AGG CTC ACC -3'

PIF3pPICZrev NotI

5'- TAA GCT GCG GCC GCC AGA ACC TCT TGG AAC AAG CGA CGA TCC ACA AAA CTG  
ATC AGA AGA CCC -3'

The PCR products were processed and ligated into pPICZ A as described.

### 3.2.20 Cloning of *PHYA* and *PHYB* from *A. thaliana* into the vector pPICZ A

The genes encoding *A. thaliana* phytochrome A and B were amplified from cDNA library like in the case of *PIF3* (3.2.19). The oligonucleotide primers used, equipped the gene with vector and gene compatible restriction sites (see below), a yeast consensus sequence for transcription initiation and the coding sequence for a C-terminal thrombin recognition sites.

*PHYA*:

AraAfwB1 (Pml I)

5'- AC GAG CAC GTG ACA ATA ATG TCA GGC TCT AGG CCG ACT CAG -3'

AraArev1 (Not I)

5'- AAG CTG CGG CCG CCA GAA CCT CTT GGA ACA AGC TTG TTT GCT GCA GCG AGT  
TCC GC -3'

*PHYB*:

AraBfwB1 (Sfi I)

5'-ACG AGG GCC CAG CCG GCC ACA ATA ATG TCA GTT TCC GGA GTC GGG GGT AGT G  
-3'

AraBrev1 (Not I)

5'- AAG CTG CGG CCG CCA GAA CCT CTT GGA ACA AGA TAT GGC ATC ATC AGC ATC  
ATG TCA C -3'

The amplified genes were processed and ligated into pPICZ A as described.

For sequencing of the vector constructs the following oligonucleotide primer were used:

*PHYA*

AthAFw600	21	ATACAAGCTCGCTGCCAAAGC
AthAFw1200	21	GGAGGTGGAACCTCGATAACCA
AthAFw1800	24	TTACTCGAAGCTAAATGATCTCAA
AthAFw2400	21	AGAAGTATTTGGGACGCAGAA
AthAFw3000	24	AATGTCTGACACTTTGTATGGAGA

*PHYB*

AthBseq435	20	CAGCCTTTCGGATGTATGAT
AthBFw600	21	GTTTTACGCCATTCTTCATAG
AthBseq1046	20	GAACCGTGTCCGAATGATAG
AthBFw1200	21	TGTTTGCCATCACACTTCTTC
AthBseq1360	20	AGGCTTTCGGTTTACAGTTA
AthBFw1800	21	AACTGCGGAAATGGATGCGAT
AthBfw1830	20	CATCCTCGTTCGTCCTTTCA
AthBFw2400	24	AATTTTTGCTGCTGACGAGAACAC
AthBFw3000	24	TTTCCTTGGAAGTGTTCATAAACGC

### 3.3 Protein Chemical Methods

#### Buffers

<b>buffer O</b>	300 mM NaCl, 10 mM Tris HCl, pH 8.0
<b>buffer A</b>	300 mM NaCl, 10% glycerol, 50 mM Tris HCl pH 8.0
<b>buffer G</b>	150 mM NaCl, 5% glycerol, 50 mM Tris HCl pH 8.0
<b>buffer P</b>	50 mM Tris HCl pH 7.8, 20 mM MgCl <sub>2</sub> , 0.1 mM Dithiothreitol, 7.5 mM Acetyl phosphate
<b>buffer U</b>	8M urea, 10 mM Tris HCl pH 8.0 at 20°C
<b>SDS sample buffer</b>	10 % SDS, 50 mM Tris HCl, 1 mM EDTA, 0.04 % bromophenolblue, 20 % 2-mercaptoethanol, pH 8.0]
<b>Towbin transfer buffer</b>	25 mM Tris, 192 mM glycine, 20% methanol <sub>tech</sub> , 0.1% SDS
<b>TPBS buffer</b>	10 mM sodium phosphate, 137 mM NaCl, 0.1% Tween 20, pH 7.6

#### 3.3.1 Test for pET28a(+) directed gene expression

An overnight *E. coli* cell culture in 50 ml LB broth was diluted 100-fold in a 5 l side baffled conical flask with terrific broth (TB) [Sambrook, et al., 1989] containing the appropriate antibiotic. This culture was then incubated on a platform shaker at 20°C and 130 rpm until the cells reached an OD<sub>600</sub> of 0.8. Expression of protein was induced by the addition of IPTG to a final concentration of 0.4 or 1.0 mM. After 2, 4 and 6 hours, a 1 ml aliquot was taken from the culture, the bacteria were collected by centrifugation in a microcentrifuge tube, and lysed by the addition of 100 µl 2x SDS sample buffer. Total cell proteins were analyzed by 2x SDS-PAGE as described under 3.3.2. A coomassie-stained prominent protein band of the appropriate molecular weight indicated the expression of the desired protein. As a reference, a sample of the bacterial culture taken prior to induction with IPTG was used. The remaining cells of the expression culture were harvested by centrifugation, resuspended in buffer A, and disrupted by sonication as described under 3.3.7. The cytosol was clarified from insoluble material by centrifugation for 10 min in a table top microcentrifuge at maximum speed.

#### 3.3.2 Standard SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Complex protein mixtures or protein preparations were separated and analyzed by standard SDS-PAGE [Laemmli, 1970]. All minigels (Hoefer) generally contained 13% acrylamide/bis-acrylamide solution (Protogel™, National Diagnostics, Atlanta, GA) in the separating gel and 7% in the stacking gel and were cast in a multiple gel caster (SE215, Mighty Small, Pharmacia Biotech). Protein samples to be analyzed were denatured by adding 2x SDS sample buffer (10 % SDS, 50 mM Tris HCl, 1 mM EDTA, 0.04 %

bromophenolblue, 20 % 2-mercaptoethanol, pH 8.0) and incubation at 80°C for 5 min. The gels were run in an SE250 Minigel Electrophoresis Unit (Pharmacia) at 130 to 180 V until the tracking dye reached the bottom of the gel. The molecular weight markers used were Prestained Protein Marker, broad range (6.5, 17, 25, 33, 48, 62, 83, 175 kD; New England Biolabs) and Kaleidoscope Prestained Standards, low (4, 8.3, 16.5, 28.1, 36.7 kD) and high range (7.7, 18.8, 33, 43.2, 91, 129, 216 kD; Bio-Rad). Proteins were visualized by staining with Coomassie brilliant blue solution (0.025% Brilliant Blue R250 (Sigma), 25% ethanol<sub>tech</sub>, 8% acetic acid) and subsequent destaining (30% methanol<sub>tech</sub>, 10% acetic acid) or transferred to nitrocellulose membrane for immuno-detection (3.3.3).

### 3.3.2.1 IEF gel electrophoresis

Purified proteins were analyzed on precast 5% polyacrylamide IEF gels (Novex) with a pH 3-7 gradient according to the manufacturers specifications. In general, 10 µl of protein solution were mixed with 20 µl Sample Buffer (Novex) and applied to the gel. The gels were run, fixed and coomassie stained as recommended by the manufacturer.

### 3.3.3 Immuno-detection of immobilized proteins (Western Blotting)

Protein samples were separated by standard SDS-PAGE (3.3.2) and subsequently transferred to a nitrocellulose membrane in a tank transfer system (Hoefer). After removing the stacking gel, the running gel was equilibrated in Towbin transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol<sub>tech</sub>, 0.1% SDS) for 5-15 minutes. The transfer sandwich was assembled under transfer buffer in the following way: the gel and the wet (H<sub>2</sub>O) nitrocellulose membrane were placed between two pieces of transfer buffer drenched blott paper (3 mm, Whatman). This sandwich was covered with a sponge on both sides and placed into a transfer cassette, taking care to remove all air bubbles from the layers. The transfer cassette was inserted into the tank filled with transfer buffer, such that the membrane was oriented towards the anode. The proteins were transferred at constant current (400 mA, approx. 100 V) within 1 hr. After disassembling the sandwich, the gel was generally stained with coomassie blue and the membrane with Ponceau S to verify transfer. For immuno-detection, the membrane was transferred into TPBS buffer (10 mM sodium phosphate, 137 mM NaCl, 0.1% Tween 20, pH 7.6) containing 10% blocking reagent (Boehringer-Roche) and incubated for 1–2 hrs at RT. The membrane was then transferred to TPBS buffer containing 5% blocking reagent and a monoclonal mouse anti-HIS antibody (Dianova) in a 1:10000 dilution. After additional 2 hrs at RT the membrane was thoroughly washed with TPBS buffer and finally incubated for 2 hrs at RT in TPBS buffer, containing 5% blocking reagent and an anti-mouse antibody horseradish peroxidase conjugate (Sigma) in a

1:2000 dilution. After washing the membrane (three times in TPBS) the ECL™ Western Blot Detection Kit was used according to the manufacturers protocol to visualize immuno-labeled proteins by exposure to an X-ray film [Kodak XOMat AR Film].

### 3.3.4 Heterologous expression of proteins in *E. coli*

A 50 ml LB *E. coli* cell culture from a single colony was grown overnight and diluted 100-fold in a 5 l side baffled conical flask with 3 l terrific broth (TB) [Sambrook, et al., 1989] containing the appropriate antibiotic. The culture was incubated on a platform shaker at 24-30°C and 130 rpm until the cell culture reached an OD<sub>600</sub> of 0.8-1.0 (10-16 hrs). The protein expression was then induced by adding IPTG to a final concentration of 1 mM and allowed to continue for 6–12 hrs. The cells were harvested by centrifugation (3000xg<sub>max</sub>, 4°C) and kept on ice through all subsequent steps. Recombinant protein was either immediately purified from the cells (3.3.7) or the cells were shock frozen in liquid nitrogen and stored at –20°C.

### 3.3.5 Heterologous expression of RcpA and RcpB

The coding sequences of the response regulators RcpA [GenBank accession number AF309559] and RcpB [AF309560] from *Calothrix sp.* PCC 7601 were amplified by PCR from genomic DNA. RcpA was subcloned into the vector pMEX8 [Medac, Hamburg, Germany], placing a His<sub>6</sub>-tag at the carboxy terminus of the protein [Hübschmann, et al., 2001]. RcpB was subcloned in to the vector pET28(a) [Novagen] between the *Nde*I and *Xho*I restriction sites, placing a thrombin recognition site and a His<sub>6</sub>-tag at the carboxy terminus. The plasmids were transformed into the *E. coli* strain B121(DE3) Gold [Novagen] by electroporation of competent cells and subsequent selection on LB agar plates containing 50 µg ml<sup>-1</sup> kanamycin. Expression was performed in TB medium [Sambrook, et al., 1989] within 6 hrs at 30°C and vigorous shaking. In all cases approximately 20-30 mg recombinant protein were obtained from 1 liter culture.

### 3.3.6 Heterologous expression of L-selenomethionine labeled RcpB

For the expression of L-selenomethionine labeled RcpB the plasmid harboring the *rcpb* gene was transformed into chemically competent cells of the methionine auxotroph *Escherichia coli* strain B834(DE3) [Novagen]. Following the manufacturers instruction, 100 nl plasmid solution were mixed with 50-80 µl competent cells (thawed on ice) in a 1.5 ml reaction tube and incubated on ice for 10 min. The cells were heat shocked at 42°C for one minute and again placed on ice for 15 min. Then, after adding 800 µl SOC medium [Sambrook, et al., 1989], the cells were transferred to a sterile cell culture tube and incubated at 37°C for one hour on a platform shaker (250-300 rpm). Selection of

positive transformants was accomplished on LB agar plates containing 50  $\mu\text{g ml}^{-1}$  kanamycin.

For expression of the *rcpb* gene, the positively transformed cells were grown in a minimal medium containing 33  $\text{mg l}^{-1}$  L-selenomethionine [Sigma] and all amino acids except L-methionine. One liter medium was prepared by combining the following:

0.5 g	Sodium chloride	
1.3 ml	Trace salts:	
	Boric acid	2.5 $\text{g l}^{-1}$
	Cobalt chloride hexahydrate	2.0 $\text{g l}^{-1}$
	Copper chloride dihydrate	1.13 $\text{g l}^{-1}$
	Manganese chloride dihydrate	9.8 $\text{g l}^{-1}$
	Sodium molybdate dihydrate ( $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ )	2.0 $\text{g l}^{-1}$
	Water to 1000 ml	
1.0 g	Citric acid monohydrate	
36 mg	Ferrous citrate, dissolved in hot conc. HCl	
4.02 g	Potassium dihydrogen phosphate	
7.82 g	Dipotassium monohydrogen phosphate trihydrate	
1.0 g	Ammonium chloride	
950 ml	Water	
	pH 7.0 with Sodium hydroxide	

After sterilization the medium was complemented by aseptically adding:

25 ml	20% Glucose, sterile filtered
0.5 ml	Thiamin, 1% aqueous solution, sterile filtered
0.5 ml	Biotin, 12 $\mu\text{g ml}^{-1}$ in 50 % 2-propanol, sterile filtered
2 ml	1 M Magnesium sulfate, sterile filtered

50 ml of the following amino acid mix (- met): sterile filtered

L-Alanine	400 $\text{mg l}^{-1}$	L-Arginine	400 $\text{mg l}^{-1}$
L-Asparagine	100 $\text{mg l}^{-1}$	L-Aspartic acid	250 $\text{mg l}^{-1}$
L-Cysteine	50 $\text{mg l}^{-1}$	L-Cystine	50 $\text{mg l}^{-1}$
L-Glutamine	400 $\text{mg l}^{-1}$	L-Glutamic acid	400 $\text{mg l}^{-1}$
Glycine	400 $\text{mg l}^{-1}$	L-Histidine	100 $\text{mg l}^{-1}$
L-Isoleucine	100 $\text{mg l}^{-1}$	L-Leucine	100 $\text{mg l}^{-1}$
L-Lysine	100 $\text{mg l}^{-1}$	L-Phenylalanine	50 $\text{mg l}^{-1}$
L-Proline	100 $\text{mg l}^{-1}$	L-Serine	400 $\text{mg l}^{-1}$
L-Threonine	100 $\text{mg l}^{-1}$	L-Tryptophane	50 $\text{mg l}^{-1}$
L-Tyrosine	100 $\text{mg l}^{-1}$	L-Valine	100 $\text{mg l}^{-1}$

0.5 ml Kanamycin, 33  $\text{g l}^{-1}$

A 5 ml LB culture from a single colony was grown at 37°C for 8 hours and then diluted in 100 ml of the above minimal medium containing 2 mg L-selenomethionine. After overnight incubation at 30°C this culture was transferred to 3 liters of minimal medium containing 33 mg ml<sup>-1</sup> L-selenomethionine and grown at 30°C, 250 rpm. When the cell density reached an OD<sub>600</sub> of 0.8–1.0 (approx. 16 h) expression was induced by adding IPTG to a final concentration of 1 mM. After 16 h at 30°C the cells were harvested by centrifugation and kept on ice during subsequent steps. The purification of SeMet-RcpB was performed as described in the following section (3.3.7). In order to protect SeMet-RcpB from oxidation, all aqueous solutions were degassed and supplemented with up to 10 mM 2-mercaptoethanol. The proteins were concentrated to 2–5 mg ml<sup>-1</sup> by centrifugation with an Amicon centrprep 10 concentrator in 150 mM NaCl, 10 mM Tris (pH 8.0). Quality and purity of the final samples were monitored by SDS-PAGE and ESI mass spectrometry. The incorporation of three SeMet residues per RcpB monomer was verified by ESI MS.

### 3.3.7 Purification of overexpressed proteins from *E. coli*

All steps were carried out on ice or in a cold room at 4°C except purification steps involving buffers containing 8M urea (room temperature). In the case where L-selenomethionine-labeled protein was purified, all buffers contained 1mM 2-mercaptoethanol and were degassed prior to use.

#### 3.3.7.1 Purification of wild-type RcpA and RcpB

Bacterial cells as obtained from 3.3.5 were resuspended in buffer A<sup>plus</sup> (buffer A (300 mM NaCl, 10% glycerol, 50 mM Tris HCl pH 8.0) supplemented with 0.5 mM EDTA, 10 mM 2-mercaptoethanol and Complete™ protease inhibitor cocktail (Roche) according to the manufacturers instructions) and disrupted by sonification. After the addition of appropriate amounts of DNase I the cytosolic proteins were separated from insoluble debris by centrifugation at 75.000xg<sub>max</sub> for 20 min. The supernatant was applied to a Ni-NTA column (1.6x10 cm) equilibrated in buffer A. The column was washed with 30 column volumes of buffer A containing 10 mM imidazole, at a flow rate of 1.5 ml min<sup>-1</sup> and developed with an imidazole gradient (1 ml min<sup>-1</sup>) from 25 mM to 250 mM imidazole in buffer A. Eluent fractions were analyzed by SDS-PAGE and those containing the protein of interest were pooled and, after concentration in a Amicon centrprep 10 device, applied (flow rate: 0.53 ml min<sup>-1</sup>) to a HiLoad 26/60 Superdex 75 prep grade column (Amersham Pharmacia Biotech) pre-equilibrated with buffer G (150 mM NaCl, 5% glycerol, 50 mM Tris HCl pH 8.0). In general, the eluted proteins were essentially pure (>95%) as judged from SDS-PAGE analysis. In cases where a thrombin cleavage site was tailored into the recombinant protein, the samples were additionally

incubated overnight at 18°C with thrombin (Pharmacia,) as described in 3.3.12 and again purified by Nickel-NTA followed by gel filtration. Proteins were concentrated to 2-5 mg ml<sup>-1</sup> by centrifugation with an Amicon centrprep 10 concentrator in 150 mM NaCl, 10 mM Tris (pH 8.0). Quality and purity of the final samples were monitored by SDS-PAGE and ESI mass spectrometry .

### 3.3.8 Purification of overexpressed proteins from *P. pastoris* and *H. polymorpha*

All steps were carried out on ice or in a cold room at 4°C. Yeast cells were resuspended in an equal volume of buffer A (3.3.7.1) supplemented with 0.5 mM EDTA and Complete™ protease inhibitor cocktail (Roche) and disrupted in a glass mill, using approximately 50 g glass beads per 20 g cells in 25 ml buffer. The glass mill was periodically cooled with a stream of supercritical liquid carbon dioxide withdrawn from a gas container equipped with a standpipe. After 2 min, disruption was completed and the cytosolic proteins were separated from insoluble debris by centrifugation at 100.000xg<sub>max</sub> for 40 min. According to the experimental design, the phytochrome proteins were either directly reconstituted with PCB in the crude lysate as described in 3.3.10 or subjected to purification as apoproteins. If the holo-protein was purified, the protein samples were illuminated with far-red light ( $\lambda > 715$  nm) and all subsequent steps were carried out under inert green light ( $\lambda_{\text{max}} = 485$  nm  $\pm$  20 nm).

The supernatant was applied to a Talon (Clontech Laboratories, Inc., Palo Alto, CA) resin filled column (1.6x10 cm) equilibrated in buffer A. The column was washed with 30 column volumes of buffer A containing 10 mM imidazole, at a flow rate of 1.0 ml min<sup>-1</sup> and developed with an imidazole gradient (1 ml min<sup>-1</sup>) from 25 mM to 250 mM imidazole in buffer A. Eluent fractions were analyzed by SDS-PAGE and those containing the protein of interest were pooled and after concentration in a Centricon K-50 device, applied to a HiLoad 26/60 Superdex 200 prep grade column (Amersham Pharmacia Biotech) pre-equilibrated with buffer G (150 mM NaCl, 5% glycerol, 50 mM Tris-HCl pH 8.0) at a constant flow rate of 0.53 ml min<sup>-1</sup>. In general, the eluted proteins were essentially pure (>95%) as judged from SDS-PAGE analysis. The quality and purity of the chromoproteins was controlled by (difference) absorption spectroscopy ( $P_r - P_{fr}$  difference spectrum).

### 3.3.9 Estimation of protein apparent molecular weights by gel filtration

To determine the apparent molecular weight of protein samples in solution, the gel filtration columns HiLoad 26/60 Superdex 75 prep grade and Superdex 200 prep grade (Amersham) were calibrated with protein standards of defined molecular weight

and conditions used throughout all purifications (3.3.7 and 3.3.8). The reference proteins were dissolved in the same buffer used for purification of expressed proteins and separated at a constant flow of  $0.53 \text{ ml min}^{-1}$ . Standards used were Ferritin (450 kD), Aldolase (158 kD), BSA (68 kD), Albumin (45 kD), Chymotrypsin (25 kD), Cytochrome C (12.5 kD) (Boehringer-Manheim). The elution volumes and the correlated molecular weights are displayed in fig. 3.1 together with the respective fitted exponential equations for the calculation of apparent molecular weights ( $y$ , kD) for a given elution volume ( $x$ , ml).

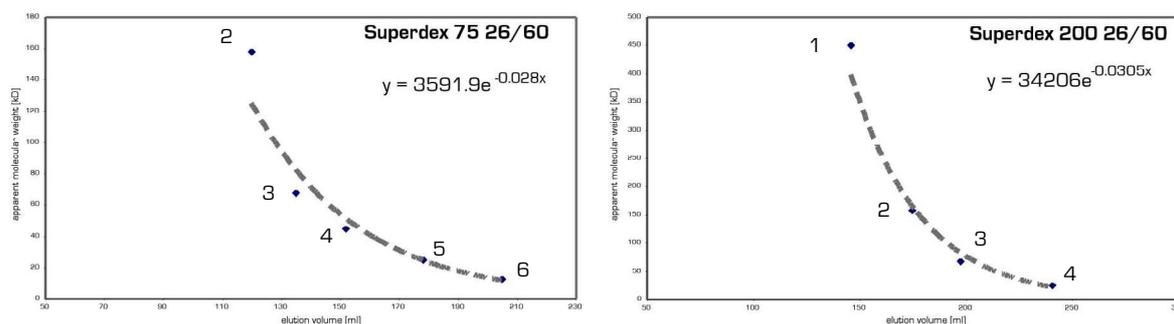


Figure 3.1. **Calibration of Size Exclusion Columns.** The obtained elution volumes for the calibration proteins were plotted against the molecular weights (kD) and an exponential curve was fitted to the resulting points (upper right corner in each diagram). (1) Ferritin (450 kD), (2) Aldolase (158 kD), (3) BSA (68 kD), (4) Albumin (45 kD), (5) Chymotrypsin (25 kD), (6) Cytochrome C (12.5 kD).

### 3.3.10 *In vitro* Reconstitution of heterologously expressed phytochromes

All steps involving Phycocyanobilin (PCB) or related chromophores were carried out under inert green light conditions ( $\lambda_{\text{max}} = 485 \pm 20 \text{ nm}$ ). PCB was isolated from the alga *Spirulina platensis* following the method from Kufer & Scheer [Kufer and Scheer, 1979a, Kufer and Scheer, 1979b]. Approximately 2 mg [the tip of a micro spatula] of PCB were dissolved in 200  $\mu\text{l}$  cold DMSO (dimethylsulfoxide) and diluted with 800  $\mu\text{l}$  water. This stock solution was used to reconstitute the apo-protein by adding aliquots to a clarified cell lysate or purified samples from 3.3.8. The formation of holo-phytochrome was monitored by following the increasing characteristic absorption at 655 nm with a UV/VIS spectrophotometer (3.5.1). To guarantee complete assembly, the chromophore was added stepwise in a 5 to 10 fold excess. After equilibration, excessive PCB was removed by gel filtration chromatography (3.3.8).

### 3.3.11 Isolation of PIF3-containing inclusion bodies from *E. coli* cells

An appropriate amount of bacterial cells (usually 20g) was resuspended in warm tap water, and 30% LDAO [N,N dimethyldodecylamine-N-oxide] in water was added to a final concentration of 0.6%. After 5 min at room temperature, lysozyme in buffer O (300 mM NaCl, 10 mM Tris HCl, pH 8.0) was added, and the cells were again kept at room temperature for additional 10 min. The cell suspension turned into a viscous paste indicating cell lysis. Chromosomal DNA was then digested by adding DNaseI and the inclusion bodies were recovered by centrifugation at 10.000xg<sub>max</sub> for 10 min. The pellet containing inclusion bodies and cellular debris was transferred to a Potter-Elvehjem homogenizer and was resuspended with buffer O. Again, the inclusion bodies were collected by centrifugation at 10.000xg<sub>max</sub> for 10 min and the homogenization/washing step was repeated twice with first buffer O supplemented with 20% glycerol, followed by one washing step with water. The such obtained inclusion bodies were shock frozen in liquid nitrogen and stored at -20°C.

### 3.3.12 Refolding of proteins from inclusion bodies

PIF3-containing inclusion bodies were purified from *E. coli* cells as described under 3.3.11. The inclusion bodies were solubilized in buffer U (8M urea, 10 mM Tris HCl, pH 8.0 at 20°C) and the solution was clarified by centrifugation (75.000xg<sub>max</sub> for 10 min). The supernatant was dialyzed overnight against buffer O supplemented with 20 mM L-arginine at 4°C. Precipitant was removed by centrifugation at 75.000xg<sub>max</sub> for 10 min and the supernatant was analyzed by SDS-PAGE.

### 3.3.13 Thrombine cleavage

To remove the terminal His<sub>6</sub>-tag from purified proteins, thrombine protease (1 U mg<sup>-1</sup> recombinant protein; Amersham Pharmacia Biotech) and CaCl<sub>2</sub> (2.5 mM final concentration) were added to the protein in buffer G (without 2-mercaptoethanol!). After incubation for 24 hrs at 18°C, the sample was applied to a Ni-NTA column (1.6x2 cm) equilibrated in buffer A with 5 mM imidazole (3.3.7) to remove residual His<sub>6</sub>-tagged proteins and poly-histidine peptides. Flow through fractions containing the protein were pooled, concentrated in a Centricon K-10 device and applied (flow rate: 0.53 ml min<sup>-1</sup>) to a HiLoad 26/60 Superdex 75 prep grade column (Amersham Pharmacia Biotech) which had been equilibrated in buffer G.

### 3.3.14 Phosphorylation assays

For the *in vitro* phosphorylation of RcpA and RcpB, 10 µl of protein solution (1 mg ml<sup>-1</sup>) in buffer G were mixed in a microcentrifuge tube with 90 µl buffer P (50 mM Tris HCl pH

7.8, 20 mM MgCl<sub>2</sub>, 0.1 mM Dithiothreitol, 7.5 mM Acetyl phosphate] and incubated at RT for 45 min. The reaction was terminated by the addition of an equal volume of 2x SDS-PAGE sample buffer supplemented with 50 mM EDTA. Phosphorylated proteins were analyzed by ESI-mass spectrometry.

## 3.4 Cloning and Expression in Yeast

Media used for cultivation of yeast strains were prepared according to the manual for cloning and expression in *P. pastoris* (Pichia Expression Kit, Invitrogen) unless otherwise indicated.

### 3.4.1 Preparation of linearized plasmid DNA for yeast transformation

Propagation of plasmids for the transformation of *P. pastoris* (pPICZ A, Invitrogen) was accomplished using the same protocol as for bacterial vectors, except that positive transformants were selected on low-salt LB agar plates containing 25  $\mu\text{g ml}^{-1}$  Zeocin (Invitrogen). The incubation time necessary for visible colonies to form was prolonged compared to the selection on ampicillin- or kanamycin-plates. After 36 to 48 hrs single colonies were picked and grown in 50 ml low-salt LB medium with 25  $\mu\text{g ml}^{-1}$  Zeocin at 37°C, 200 rpm overnight. The plasmids were isolated as described in 3.2.2. Linearization of plasmid DNA was accomplished by incubation with the appropriate restriction enzyme in microcentrifuge tubes. The standard reaction contained up to 10  $\mu\text{g}$  of DNA and 1 unit restriction enzyme per  $\mu\text{g}$  DNA in a final volume of 100  $\mu\text{l}$  aqueous buffer (supplied with the enzyme). Reactions were incubated at 37°C for 45-60 min and monitored by agarose gel electrophoresis, to ensure complete digestion. The linearized plasmid DNA was purified as described previously and stored as a buffered aqueous solution (EB buffer, Qiagen) at -20°C.

### 3.4.2 Preparation of electrocompetent *P. pastoris* cells

According to the manufacturer's protocol, electrocompetent yeast cells were prepared as follows. A single colony of the *P. pastoris* strain GS115 was grown overnight in 5 ml YPD in a 50 ml conical flask on a platform shaker at 30°C, 250 rpm. From that, 500 ml of fresh medium in a 2 l flask were inoculated with 0.1–0.5 ml and again incubated at 30°C, 250 rpm overnight to a final  $\text{OD}_{600}$  of 1.3-1.5. The cells were harvested by centrifugation (1500 $\times g_{\text{max}}$ , 4°C, 5 min) and kept on ice during the following steps. The pellet was resuspended with 500 ml of ice-cold sterile water and the cells were centrifuged as before. The pellet was then resuspended with 250 ml of ice-cold sterile water and again centrifuged. The washed cells were resuspended in 20 ml of ice-cold 1M sorbitol, centrifuged, finally resuspended with 1-2 ml ice-cold 1M sorbitol and dispensed as 80  $\mu\text{l}$  aliquots into microcentrifuge tubes. Electrocompetent cells were generally subjected to electroporation on the same day, in order to guarantee a high transformation efficiency.

### 3.4.3 Transformation of *P. pastoris* cells

For the transformation of *P. pastoris*, concentrated linearized plasmid DNA was prepared as described above. 80  $\mu$ l of electrocompetent cells were mixed with 10  $\mu$ l of the DNA solution (10  $\mu$ g in EB buffer) and transferred to an ice-cold 0.2 cm electroporation cuvette. After incubating on ice for 5 min, the cells were pulsed using a GenePulser (Bio-Rad) electroporation device with the parameters for yeast [*Saccharomyces cerevisiae*: charging voltage 1500 V, capacitance 25  $\mu$ F, resistance 200  $\Omega$ ]. Generally a pulse length of 5-10 ms was generated with 0.2 cm cuvettes. After pulsing, 1 ml of ice-cold 1M sorbitol was added, the resuspended cells transferred to a sterile 15 ml culture tube and incubated at 30°C for one hour. Then 1 ml of YPD medium was added and the cultures were shaken at 30°C, 200 rpm for additional two hours. All of the cells were plated out by spreading 300  $\mu$ l on 150 mm plates, containing MDH-Agar with increasing concentrations of Zeocin™ (Invitrogen), generally 100, 200 and 500  $\mu$ g ml<sup>-1</sup>. The plates were incubated at 30°C for 2-4 days until colonies formed.

### 3.4.4 Control of gene expression in recombinant *P. pastoris* strains

Genes cloned into the vectors pHIL D2 or pPICZ A were expressed under the control of the AOX1 promotor and thus induced and regulated by the presence of methanol as the sole carbon source. In general, those colonies that had grown on plates with the highest Zeocin™ concentration were tested for expression of the introduced gene. Single colonies were picked and grown in 250 ml conicals containing 50 ml MGYH medium, at 30°C, 250 rpm overnight to a final OD<sub>600</sub> of approx. 4-6. To induce protein expression, the cells were harvested by centrifugation (3000xg<sub>max</sub>, 5 min), transferred to 250 ml of fresh MMH medium in 1000 ml conicals and incubated at 30°C, 250 rpm. At several time points (0, 6, 12, 18, 24, 36 hours) 1ml aliquots of the expression cultures were transferred to 1.5 ml microcentrifuge tubes, centrifuged (14 000xg<sub>max</sub>, 5 min) and the cells were frozen in liquid nitrogen until further analysis. Levels of expressed recombinant protein were analyzed by denaturing SDS-PAGE (see 3.3.2) and immuno-detection with HIS-tag specific antibodies (Western-Analysis, see 3.3.3), and/or, in the case of tetrapyrrole binding proteins (phytochromes and phytochrome fragments), by in vitro reconstitution with PCB and UV/VIS spectroscopical characterization (see 3.5.1). For SDS PAGE and immuno-detection, pelleted wet cells were resuspended in an equal volume of Yeast sample buffer (6 M urea, 60 mM Tris HCl, pH 6.8, 2% SDS, 5% 2 mercaptoethanol, 10% glycerol, Bromophenol Blue) and denatured at 95°C for 5 min. After centrifugation for 10 min at 21 000 g<sub>max</sub> in a table top centrifuge (Eppendorf) the protein rich cell extracts were separated on SDS gels as described.

### 3.4.5 Fermentation of *P. pastoris* and *H. polymorpha*

The fermenter was autoclaved twice, first charged with water, than charged with the fermentation medium as described below.

#### Media

##### Fermentation Basal Salts Medium

For 1 liter, the following ingredients were combined:

Phosphoric acid, 85%	13.2 g
Calcium sulfate	0.93 g
Potassium sulfate	18.2 g
Magnesium sulfate heptahydrate	14.9 g
Potassium hydroxide	4.13 g
Glycerol	40.0 g
Ammonium hydroxide 33%	pH to 5.0

Water to 1 liter

Added to fermenter and sterilized; pH adjusted to 5.0 with ammonia.

##### PTM1 Trace Salts

Cupric sulfate-5 H <sub>2</sub> O	6.0 g
Sodium iodide	0.08 g
Manganese sulfate-H <sub>2</sub> O	3.0 g
Sodium molybdate-2 H <sub>2</sub> O	0.2 g
Boric Acid	0.02 g
Cobalt chloride	0.5 g
Zinc chloride	20.0 g
Ferrous sulfate-7 H <sub>2</sub> O	65.0 g
Biotin	0.2 g
Sulfuric Acid	5.0 ml

Water to a final volume of 1 liter

filter sterilized and stored at room temperature.

##### Vitamins

Biotin	6mg dissolved in 10 ml 50% iso-propyl alcohol
Thiamin	2.0 g dissolved in 90 ml water

Both solutions were mixed and filter sterilized; stored at 4°C.

##### Glycerol Nutrient

50 % Glycerol containing 12 ml PTM1 Trace Salts per liter

filter sterilized

##### Methanol Nutrient (induction of the AOX-Promotor)

100% Methanol containing 12 ml PTM1 Trace Salts per liter

filter sterilized

### Base

Ammonium hydroxide 33%

### Fermentation procedure:

A single colony of the desired yeast strain was grown overnight in a 5ml YPD flask culture (OD<sub>600</sub> = 2-6) at 37°C, 250-300 rpm. 50 ml fermentation medium (containing PTM, see below) were inoculated with 1 ml from the overnight culture and grown overnight at 30°C, 250-300rpm.

The sterilized fermenter containing an appropriate volume of Basal Salts Medium, (supplemented with 4.35 ml PTM1 Trace Salts per liter and Vitamins (10 ml per liter, pH adjusted to 5.0)) was equilibrated to 30°C and agitation/aeration were set to maximum (1000 rpm, maximum air). Inoculation of the fermenter was done aseptically with approx. 1- 5% initial fermentation volume from the overnight culture. This batch culture was grown until glycerol was completely metabolized (24–36h, indicated by a static OD<sub>600</sub> and a stable pH). The OD<sub>600</sub> has usually reached 60-80 at this point. Correct pH (5.0-5.5) of the culture was controlled *via* the automated feeding of Base solution.

Induction of expression was initiated by feeding Methanol Nutrient at an initial rate of 3-4 ml hr<sup>-1</sup> l<sup>-1</sup> and cooling the fermenter to 27°C. After 2 hrs the feed rate was slowly increased within 6 hrs to a final rate of 10 ml hr<sup>-1</sup> l<sup>-1</sup>. This rate was usually maintained throughout the remainder of the fermentation. The duration of the entire methanol feed-batch phase depended on the specific experimental setup and the cellular mass to be obtained, and was in the range of 12 to 36 hrs.

The cells were harvested by centrifugation (3000xg<sub>max</sub>, 10 min), shock frozen in liquid nitrogen and stored at -20°C.

### 3.4.5.1 Fermentation of yeasts for expression of <sup>15</sup>N-labeled proteins

The following protocol has been used for the fermentation of the methylotrophic yeasts *Hansenula polymorpha* and *Pichia pastoris* in the presence of <sup>15</sup>N-ammonium chloride as the sole nitrogen source. <sup>15</sup>N-Ammonium chloride was purchased from Cambridge Isotope Laboratories, Inc. (Andover, MA) and contained >95% <sup>15</sup>N-isotope.

#### Media

##### Fermentation Basal Salts Medium

For 1 liter, the following ingredients were combined:

Phosphoric acid, 85%	13.2 g
<sup>15</sup> N-Ammonium chloride	6.15 g
Calcium sulfate	0.93 g
Potassium sulfate	18.2 g
Magnesium sulfate heptahydrate	14.9 g
Potassium hydroxide	4.13 g
Sodium hydroxide	0.30 g
Glycerol	40.0 g

Water to 1 liter

Add to fermenter and sterilize. Adjust pH to 5.0 with KOH.

##### PTM1 Trace Salts

Cupric sulfate-5 H <sub>2</sub> O	6.0 g
Sodium iodide	0.08 g
Manganese sulfate-H <sub>2</sub> O	3.0 g
Sodium molybdate-2 H <sub>2</sub> O	0.2 g
Boric Acid	0.02 g
Cobalt chloride	0.5 g
Zinc chloride	20.0 g
Ferrous sulfate-7 H <sub>2</sub> O	65.0 g
Biotin	0.2 g
Sulfuric Acid	5.0 ml

Water to a final volume of 1 liter

filter sterilized and stored at room temperature.

##### Glycerol Nutrient

50 % Glycerol containing 12 ml PTM1 Trace Salts per liter

filter sterilized

##### Methanol Nutrient

100% Methanol containing 12 ml PTM1 Trace Salts per liter

filter sterilized

### Base

1-2 M KOH containing 5% w/v <sup>15</sup>N-Ammonium chloride  
(kept on ice to prevent <sup>15</sup>N-Ammonia from evaporating)

### Fermentation procedure

Inoculum. 5ml of a flask culture were inoculated with a single colony of the desired yeast strain and grown overnight in YPD at 37°C, 250-300rpm to an OD<sub>600</sub> of 2-6. 50 ml of the fermentation medium (supplemented with 4.35 ml PTM1 Trace Salts per liter, see below) were inoculated with 1 ml of the YPD culture and grown overnight at 30°C, 250-300rpm.

Initial glycerol feed phase. The fermenter was charged with Fermentation Basal Salts Medium, autoclaved and the pH adjusted to 5.0, aseptically, with 5 M KOH. 4.35 ml of the above PTM1 Trace Salt solution per liter fermentation volume were added aseptically *via* a sterile syringe and a septum in the fermenter lid. The fermenter was then prepared by adjusting the temperature of the medium to 30°C and setting agitation and aeration to maximum (usually 1000 rpm and maximum air flow). After equilibration, the fermenter was inoculated with approx. 1- 5% initial fermentation volume of the overnight culture. This batch culture was grown until the glycerol was completely consumed (24–36h, indicated by a constant OD<sub>600</sub>). The OD<sub>600</sub> generally has reached 60-80 at this point. The following Glycerol feed-batch phase was omitted whenever the hitherto obtained cell density was sufficiently high.

### Glycerol feed-batch phase

The glycerol feed-batch phase was initiated by feeding 50% glycerol nutrient at 18.5 ml hr<sup>-1</sup> l<sup>-1</sup>. This should be carried out for about 4 to 6 hrs depending on the cellular yield to be achieved (generally the OD<sub>600</sub> reached 150-200). After termination of the glycerol feed, the cells were grown for another 6-12 hrs to ensure that all glycerol was completely metabolized (glycerol depresses the AOX1-promotor).

### Methanol induction

The induction of the AOX1 promotor was accomplished by feeding Methanol Nutrient at an initial rate of 3 ml hr<sup>-1</sup> l<sup>-1</sup> and after cooling the fermenter to 27°C. After 2 hrs the feed rate was increased slowly within 6 hrs to a final rate of 10 ml hr<sup>-1</sup> l<sup>-1</sup>. This rate was maintained throughout the remainder of the fermentation procedure. Samples were taken at distinct time points for control of expression. The duration of the entire methanol feed-batch phase depended on the specific experimental setup. It was routinely in the range of 12 to 24 hrs.

## 3.5 Spectroscopic Methods

### 3.5.1 UV/VIS Spectroscopy of recombinant phytochromes

For the spectroscopic characterization and the recording of  $P_r/P_{fr}$  difference spectra, a Shimadzu UV-2401 PC Spectrophotometer and quartz cuvettes (path length 10 mm) were used. Samples were generally illuminated for 30-120 sec (concentration dependent) with red light ( $635 \pm 7$  nm, interference filter) or far-red light ( $>715$  nm, cutoff filter) in order to generate the two photoisomeric states  $P_r$  and  $P_{fr}$ , respectively. The spectra were recorded in the range of 200 to 800 nm and difference spectra were calculated by subtracting the  $P_{fr}$ -spectrum from the  $P_r$ -spectrum.

### 3.5.2 Circular dichroism spectroscopy

Circular dichroism (CD) is a sensitive technique for the determination of secondary structure elements in proteins and for monitoring structural changes in macromolecules. The secondary structural elements in proteins show specific bands in the far UV CD spectrum and allow, on a purely empirical basis, the correlation of a protein's CD spectral characteristics with its secondary structure composition. A more detailed discussion on this relationship is given in several reviews [Adler, et al., 1973, Bazzi and Woody, 1985, Johnson, 1985, Johnson, 1988, Johnson, 1990, Sreerama, 1993, Yang, et al., 1986].

The result of a CD measurement is the differential absorption of circular polarized light as a function of the wavelength and is represented as the ellipticity  $\Theta$  in radians or degree:

$$\Theta_r = \frac{2.303}{4} \cdot (A_L - A_R) \cdot [rad]$$

$$\Theta_d = \frac{2.303}{4} \cdot (A_L \cdot A_R) \cdot \frac{180}{\pi} \cdot [deg]$$

To be able to compare the ellipticity of different probes, the values are normalized:

$$\Theta = \frac{\Theta_d \cdot M}{c \cdot l} \quad [\Theta] = \frac{deg \cdot cm^2}{mol}$$

$$\Theta_{mr} = \Theta_d \cdot \frac{M}{c \cdot l \cdot n_r}$$

with	$A_L/A_R$	absorbance of the left/right handed component of the incident linear polarized light
	$\Theta_r$	ellipticity in [ <i>rad</i> ]
	$\Theta_s$	ellipticity in [ <i>deg</i> ]
	$\Theta$	molar ellipticity
	$\Theta_{mr}$	mean molar ellipticity per residue
	$M$	molecular weight
	$c$	concentration
	$l$	path length
	$n_r$	number of residues

Protein samples as obtained from 3.3.8 were dialyzed overnight against buffer C (100 mM NaCl, 10 mM Tris HCl pH 8.0, 1 mM *rac*-DTT), to remove traces of 2-mercaptoethanol, concentrated (Centricon K-50) to an absorption  $A_{660}$  of 1 ( $A_{654}$  of 1.1), corresponding to 10  $\mu$ M, and subjected to CD spectroscopy (JASCO J-715 Spectropolarimeter).

CD spectra were recorded at 20°C, using a 2 mm path length quartz cuvette for measurements in the far-UV region (185-250 nm) and a 10 mm path length cuvette in the ultraviolet and visible region (250-800 nm). For the measurement of the  $P_r$ - or  $P_{fr}$ -states, the samples were irradiated with red light (635  $\pm$  7 nm, interference filter) or with far red light > 715 nm (cut-off filter) for ca. 30 seconds. For each CD spectrum, ten single recordings were averaged; after each recording, the samples were irradiated with the primarily applied light quality in order to regenerate the  $P_r$  - or  $P_{fr}$  -state.

Secondary structure analysis based on the CD data was done with the program K2D (Andrade et al., 1993). In principle, a CD spectrum of a protein  $P(\lambda)$  can be regarded as a linear combination of  $n$  basis spectra  $B_n(\lambda)$ .

$$S(\lambda) = \sum_{n=1}^N f_n B_n(\lambda)$$

where  $f_n$  is the fraction of the  $n$ -th secondary structure. Thus, with basis spectra of protein with known secondary structure, the secondary structure composition of any protein can be approximated.

### 3.5.3 $^1\text{H}$ - $^{15}\text{N}$ HSQC Nuclear Magnetic Resonance Spectroscopy

The  $^1\text{H}$ - $^{15}\text{N}$  HSQC (or  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single quantum coherence) NMR experiment is a two dimensional heteronuclear NMR experiment [Bodenhausen, 1980] in which each amino acid residue (except for proline) in a given protein is represented by one peak in the spectrum. The chemical shift of a peak at  $(\omega_1, \omega_2)$ , where  $\omega_1$  and  $\omega_2$  are the amide  $^{15}\text{N}$  and  $^1\text{H}$  shifts, respectively, varies with the chemical surrounding of the amino acid amide group in the protein. Side chains are not represented in the spectrum, apart from the side chain amides of Gln and Asn. A specifically folded protein structure will generally produce a  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum with a broad distribution of well separated signals. It is therefore suited for the control of a protein sample with regard to its purity, folding state, stability etc.

#### 3.5.3.1 $^1\text{H}$ - $^{15}\text{N}$ HSQC spectra of 59 kD phytochrome A

The uniform labeling of the protein with  $^{15}\text{N}$  is an essential prerequisite, and, in the case of recombinant phytochrome, was realized as described in 3.4.5. NMR spectra were recorded on a Bruker Avance DMX 750 MHz multinuclear NMR spectrometer at the Bavarian NMR Center [Technische Hochschule München, Garching]. The samples were prepared as described under 3.3.8 and checked for intact photoreversibility prior to the experiment. The protein concentration was elevated (3.3.8) to at least 100  $\mu\text{M}$  or higher in 300 mM NaCl, 10 mM Tris HCl pH 8.0, 10 mM 2-mercaptoethanol. The samples were irradiated with either red ( $635 \pm 8$  nm, interference filter, Schott) or far red light ( $> 715$  nm, cut-off filter, Schott) for 120 seconds in the NMR tube and subsequently positioned in the NMR spectrometer's magnet. Recording of the HSQC-spectra was performed over-night at room temperature.

## 3.6 Protein Crystallographic Methods

The word 'crystal' is derived from the Greek root 'krystallos' meaning clear ice. Like ice, crystals are chemically well defined, and many among them are of transparent and glittering appearance and they have, at all times, attracted mankind's attention due to their magic beauty. Their flat faces and their physical anisotropy reflect the regular packing of molecules, atoms or ions in the crystal. Even though obtaining crystals of high quality that are suitable for X-ray structure determination is the compulsive basis, it is also the least understood step in the process of crystal structure analysis. Biocrystallization, the crystallization of biological macromolecules like proteins, DNA, RNA or their complexes, like any crystallization, is a multiparametric process involving the three classical steps of nucleation, growth, and cessation of growth. Usually, a solution containing the molecule to be crystallized must initially be brought to a state of supersaturation, which will ideally force the macromolecules to arrange in a regular 3-dimensional manner. Because of the complexity of the particles (proteins, nucleic acids etc.) involved in biocrystallization, the crystallization itself is often the primary obstacle that sometimes proves to be insurmountable.

### 3.6.1 Protein crystallization techniques

All preliminary crystallization experiments were performed using the hanging or sitting drop vapor diffusion method [Davis & Segal, 1971] and a *sparse matrix* strategy [Jancarik and Kim, 1991]. The latter was chosen because of the commercial availability of standardized kits covering a broad factorial space of crystallization conditions and providing the benefits of industry standards, meeting the high demands regarding purity and consistency over different charges. The *sparse matrix* strategy is an empirical approach, based on crystallization conditions that have successfully been utilized to crystallize biological macromolecules as introduced by Jancarik and Kim, 1991.

For the initial screening for conditions compatible with a given protein, a standard set of 50-150 crystallization conditions was tested and evaluated by judging the crystallization drops under a microscope. The compositions of the standard screening conditions is listed in the appendix B. These conditions vary several parameters such as pH-value, type of buffer, type and concentration of precipitant and electrolyte composition in a non-systematic manner. Conditions that immediately or within the first days produced brownish cloudy precipitate were excluded from further testing. Conditions that yielded any kind of crystalline arrangement in the droplets served as starting points for the optimization of the respective condition. This was accomplished by designing two dimensional grids with two varying conditions, i.e. pH-value and precipitant and thereby narrowing the variation space of the crystallization parameters involved.

In the case of experiments involving phytochrome receptor proteins, all crystallization screens were set up under well defined light conditions, either inert green light ( $\lambda_{\max} = 485$  nm) or far red light (cutoff  $\lambda > 715$  nm). Prior to incubation in the dark at 18°C, the droplets were additionally exposed to far red or red ( $\lambda_{\max} = 635 \pm 7$  nm) light in order to shift the photo equilibrium to either the  $P_r$  or the  $P_{fr}$  form.

### 3.6.2 Crystallization of RcpA

Crystals were obtained at room temperature with the hanging drop vapor diffusion method using a well solution that contained 1.4 M sodium citrate, 150 mM NaCl, 10% glycerol and 100 mM HEPES, pH 7.0. The crystallization droplets contained the RcpA solution as described under 3.3.7.1 mixed with an equal volume of the well solution. Crystals appeared after two weeks and grew to a size of approx. 100x100x30 microns. RcpA crystallized in space group  $P2_12_12$  with unit cell dimensions  $a = 76.7$  Å,  $b = 85.2$  Å,  $c = 44.1$  Å and two molecules in the asymmetric unit [AU].

### 3.6.3 Crystallization of RcpB and SeMet-RcpB

Crystals grew at 4°C in hanging drops by vapor diffusion against a reservoir of 3.75 M sodium formiate, 5% glycerol and 1 mM 2-mercaptoethanol. The droplets contained equal volumes of reservoir solution and protein solution, the latter containing 2-5 mg ml<sup>-1</sup> RcpB in: 150 mM sodium chloride, 10 mM Tris HCl pH 8.0, 10% glycerol, 1mM 2-mercaptoethanol. Crystals appeared within several days and reached a final size of approx. 200 microns in each dimension after one week. Both, RcpB and SeMet-RcpB formed crystals belonging to the space group  $P4_12_12$  with the unit cell dimensions  $a = b = 72.2$  Å,  $c = 142.8$  Å and two molecules per AU (solvent content 50%).

### 3.6.4 X-Ray structural analysis

A detailed description of the complex physical principles of X-ray crystallography is beyond the scope of this work. For introductory and in-depth treatises on the theoretical background as well as on the theory of the applied methods, the following monographs are advisory: [Blundell and Johnson, 1976, Drenth, 1999, Giacovazzo, et al., 1992, Hahn, 1995, McPhearson, 1999, Stout and Jensen, 1989].

### 3.6.5 Data collection

All diffraction data sets were collected using the oscillation method [Arndt and Wonacott, 1978] under cryo-conditions at 100 or 110 K. In general, the crystals were mounted on cryo loops [Hampton Research, Laguna Niguel, CA] and attached to the

magnetic base of the goniometer head of the X-ray diffractometer. Once mounted, the crystals were flash-frozen by immersing them into a cold nitrogen gas stream from a cold nozzle mounted on the X-ray instrument. The crystals were rotated (oscillated) through small angles of 0.5–1.0°, depending on the mosaic spread of a given crystal and the signal to noise ratio, about an axis perpendicular to the X-ray beam. To ensure the collection of a complete dataset in dependency of the crystallographic symmetry and the crystal orientation, the minimal rotation range required was determined by using the STRATEGY option of the MOSFLM software package (Leslie, 1992).

### 3.6.5.1 RcpA native data collection

A high resolution native data set was collected to 1.9 Å at 100 K at the Max-Planck *Wiggler-Beamline* BW6 (DESY, Hamburg) using a MAR CCD detector (MAR Research, Hamburg). The wavelength of the X-ray beam was 1.05 Å.

### 3.6.5.2 RcpB data collection

An RcpB native data set was collected to 1.75 Å at 100 K at the Max-Planck *Wiggler-Beamline* BW6 (DESY, Hamburg) using a MAR CCD detector (MAR Research, Hamburg). The wavelength of the X-ray beam was 1.05 Å.

Heavy atom derivatives of RcpB were prepared as outlined in 3.6.8.3 and by using different heavy atom salts (HgCl<sub>2</sub>, MeHgCl, Cr, GdCl<sub>2</sub>, K[AuCl<sub>4</sub>], Pb(OAc)<sub>2</sub>, K<sub>2</sub>[PtCl<sub>6</sub>]) in concentrations from 1 to 100 mM. Data sets of putatively derivatized crystals were collected using a Rigaku RU-200 CuK<sub>α</sub> rotating anode X-ray generator (Rigaku, Tokyo, Japan; CuK<sub>α</sub>: λ=1,5418 Å) and an imaging plate detector (Mar Research, Hamburg). The maximum resolutions to which the data were collected depended on a given crystal and were generally in the range of 1.9-2.5 Å.

A four-wavelength, 1.6 Å multiwavelength anomalous dispersion (MAD) data set of the SeMet-RcpB derivative was collected from a single cryo-cooled crystal at beamline BW6 as summarized in the following Table.

Data set	Wavelength λ	Rot. segments	Rotation width	max. resolution
Edge	0.9811 Å	51-111°, 231-291°	0.5°	1.60 Å
Peak	0.9800 Å	51-111°, 231-291°	0.5°	1.60 Å
Remote	0.9500 Å	51-111°, 231-291°	0.5°	1.60 Å
Native	1.05	51-111°	0.25	1,36 Å

### 3.6.6 Data processing and reduction

Data processing and reduction was accomplished using the software packages HKL [Otwinowski and Minor, 1997], MOSFLM [Leslie, 1992] and CCP4 [Collaborative Computational Project, 1994]. The preliminary determination of the crystal orientation with respect to the laboratory system (as defined by the arrangement of X-ray source and detector), the space group and the unit cell dimensions were done with either DENZO [HKL package] or MOSFLM. During autoindexing, the cell parameters and crystal orientation as well as the film orientation and the detector distance were further refined. From autoindexing the intensities of the integrated and indexed reflections were obtained and further subjected to SCALEPACK [HKL package] in order to scale data from subsequent images. A measure for the quality of a data set can be estimated by comparing the structure factor amplitudes  $|F(hkl)|$  for symmetry related reflections and is expressed as the  $R_{merge}$  value:

$$R_{merge}(F) = \frac{\sum_{hkl} \sum_i |F_i(hkl)| - \overline{|F(hkl)|}}{\sum_{hkl} \sum_i |F_i(hkl)|}$$

$R_{merge}$  for  $i$  observations each of  $n$  independent reflections.  $\overline{|F(hkl)|}$  is the average value for the structure factor amplitude of the  $i$ -th observations of a given reflection.

The actual values of the structure factor amplitudes can be approximated by the square root of the measured and scaled intensities:

$$|F(hkl)| = \sqrt{I(hkl)}$$

By applying Bayesian statistics, the experimental data can be further corrected to match a theoretical Wilson distribution [Wilson, 1949]. Thereby, otherwise negative reflection intensities (with respect to the background), and weak reflections can be treated correctly. This is done by TRUNCATE [French, 1978].

### 3.6.7 Estimation of the number of molecules per unit cell

The number of molecules per unit cell for a given crystal can be estimated by the method proposed by Matthews [Matthews, 1968]. According to this, the ratio of the unit cell volume and the molecular weight is in the range of 1.7-3.5 Å<sup>3</sup> Da<sup>-1</sup> with most values

around  $2.15 \text{ \AA}^3 \text{ Da}^{-1}$ . This empirical number is called Mathews' coefficient  $V_M$  and is given by the following relations:

$$V_M = 1.23/V_{protein}$$

$$V_{protein} = \frac{V_{spec}}{V_M \times N} = \frac{V_{protein/cell}}{V_{cell}}$$

The specific volume of a protein  $V_{spec}$  is always approximately  $0.74 \text{ cm}^3 \text{ g}^{-1}$ ;  $N$  is Avogadro's number.

### 3.6.8 Solution of the phase problem

The underlying relationship between the distribution of electrons in a unit cell (electron density) and the set of structure factors for a given diffracting crystal, is given by a Fourier transformation:

$$\rho(x,y,z) = \frac{1}{V} \sum |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)]$$

The amplitudes  $|F(hkl)|$  can be determined from measured intensities and if the phases  $\alpha(hkl)$  were known too, the electron densities could be calculated for all points  $(x,y,z)$  in the unit cell. Unfortunately though, the phase information associated with each reflected beam can not be retrieved directly from the X-ray experiment, and the problem of phase determination indeed represents the fundamental problem of macromolecular crystallography. In principle there are four ways to solve the phase problem.

1. Patterson search or molecular replacement (MR)
2. single or multiple isomorphous replacement (SIR, MIR)
3. multiple wavelength anomalous dispersion (MAD)
4. direct methods, so far restricted to small molecule X-ray crystallography

#### 3.6.8.1 The Patterson search

The fastest method is the so called Patterson search or molecular replacement method [Hoppe, 1957, Rossmann, 1972, Rossmann, 1962], for which the similarity of the unknown structure to an already known structure is a prerequisite. If such a model

exists, a vector map (Patterson map) of both, the model structure and the unknown structure can be calculated by the Patterson function (Patterson, 1934).

$$P(uvw) = \frac{1}{V} \sum_{hkl} |F(hkl)|^2 \cos[2\pi(hu + kv + lw)]$$

This function represents a Fourier summation with measured intensities  $|F(hkl)|^2$  as coefficients and with all phase angles equal to zero. The Patterson map of a protein structure represents all interatomic distances in the real structure as vectors from the origin to maxima in the Patterson map. If the Patterson function of the known structure can be superimposed on the protein for which the structure is not yet known, a preliminary estimation of the phases can be derived. For structurally homologous proteins the Patterson function is very similar and the self-Patterson vectors (vectors between atoms belonging to the same molecule) can supply the transformational relationship between the known and the unknown molecular structure. In order to transfer the known protein molecular structure from its crystalline arrangement to the crystal of the unknown protein, rotation and translation searches have to be performed, which will ideally result in a transformation matrix relating the two molecules.

### The Rotation Function

The rotation search is a real space three dimensional search. The Patterson function of the model structure is rotated  $P_r(u_r)$  and superimposed on the Patterson function of the unknown structure  $P(u)$ . If the rotation is correct, maximum overlap  $R(\alpha, \beta, \gamma)$  between the two maps will occur, which can be monitored by a high correlation coefficient  $C$  and low  $R$ -factor, see next paragraph "Translation Function".

$$R(\alpha, \beta, \gamma) = \int_U P(u) \times P_r(u_r) du$$

$U$  is the volume of the Patterson map where the self-Patterson peaks are located.  $R$  depends of the rotation angles given by a rotation matrix  $[C]$  which transforms a vektor  $x$  to a rotated vektor  $x_r$ .

$$x_r = [C] x$$

### The Translation Function

The translation search is not incorporated in the rotation search but is instead performed separately in order to reduce the complexity of the MR problem. A combined

translation /rotation search would represent a six-dimensional problem which is still a hard task to perform, even for fast computers. The translation search is carried out after the correct rotational orientation has been found. In principle, the translation search is a process of trial and error. The known molecule is moved through the asymmetric unit, and structure factors  $F_{calc}$  are calculated and compared with the observed structure factors  $F_{obs}$  by calculating a correlation coefficient  $C$  and an  $R$ -factor as a function of the molecular position.

$$C = \frac{\left| \sum_{hkl} \left( |F_{obs}|^2 - \overline{|F_{obs}|^2} \right) \times \left( |F_{calc}|^2 - \overline{|F_{calc}|^2} \right) \right|}{\left[ \sum_{hkl} \left( |F_{obs}|^2 - \overline{|F_{obs}|^2} \right)^2 \sum_{hkl} \left( |F_{calc}|^2 - \overline{|F_{calc}|^2} \right)^2 \right]^{1/2}}$$

$$R = \frac{\sum_{hkl} \left\| |F_{obs}| - k |F_{calc}| \right\|}{\sum_{hkl} |F_{obs}|}$$

### 3.6.8.2 Solution of the phase problem for RcpA by MR

The initial phases for the native data set of RcpA were determined by molecular replacement (MR) using the structure of SeMet-RcpB as the search model. The latter was solved by MAD techniques as described below (3.6.8.4). The number of molecules in the unit cell (and in the asymmetric unit) was estimated prior to the MR calculations by the method described in 3.6.7. A two molecule solution was found with AMoRe (Navaza, 1994) giving a correlation coefficient of 55% and an  $R$ -factor of 33%.

### 3.6.8.3 The Multiple Isomorphous Replacement method (MIR)

This method, which was first applied by Perutz (Perutz, 1956), determines phase angles of a molecular structure in principle by the difference in X-ray intensities of related reflections of two data sets, induced by the presence of strong atomic scatterers (heavy atoms) in one of the data sets. Therefore, MIR requires the isomorphous attachment of heavy atoms to specific sites in the protein and subsequent identification of the coordinates of these heavy atoms in the unit cell.

The introduction of heavy atom substances is accomplished by soaking of the crystals in a solution equal to the mother liquor with the respective heavy atom compound added in varying concentrations. In general, the crystals were transferred to the soak solution

with a cryo-loop and allowed to equilibrate for 1-16 hrs. Subsequently each crystal was immediately mounted to the X-ray diffractometer and, provided that isomorphy was preserved, a complete data set was measured. To judge whether a measured crystal was a derivative or not, the data was processed and scaled to the native data set with the software SCALEIT (CCP4) and the isomorphous differences were calculated (TRUNCATE).

The subsequent steps involve the calculation of difference Patterson maps and the interpretation of peaks in specific Harker sections (special planes in the unit cell) to derive the positions of the heavy atoms. Difference Fourier analysis and improvement of preliminary phases will eventually lead to interpretable electron density maps.

#### 3.6.8.4 The Multiple Anomalous Wavelength Dispersion method (MAD)

Electrons in an atom can be considered as oscillators with natural frequencies (distinct energies). If the frequency of the incident X-ray beam is near to a natural frequency of an electron in the atom, resonance will take place rather than elastic scattering and the absorbed photon will be re-emitted with an altered phase angle, an effect referred to as anomalous scattering. When a protein contains anomalous scatterers in its molecules, Friedel's law breaks down and a structure factor  $F(hkl)$  is no longer equal to its Friedel's mate  $F(\bar{h}\bar{k}\bar{l})$ . This effect strongly depends on the wavelength of the initial X-ray beam. Elements of the first periods, H, C, N, O, S, P, usually building up biological macromolecules, are no suitable anomalous scatterers and so, heavier atoms, such as selenium, have to be introduced into the protein molecule. This is basically accomplished by growing a methionine auxotroph organism on a selenomethionine substrate (see 3.3.6). In heavy atoms, the energy of transitions between electronic ground state levels and excited states are near the energy of photons used in X-ray diffraction. As mentioned, the anomalous effect depends on the wavelength of the incident photon (dispersion effect) and has its strongest impact at one of the atom's X-ray absorption edges. Phase information of a scattered beam can be derived from the resonance effect by correcting the expression for the atomic scattering factor  $f$  with a complex expression for the anomalous scattering (Hendrickson, et al., 1985, Karle, 1980):

$$f_{anom.} = f_0 + \Delta f + if''$$

with  $\Delta f$  being the change in the electron scattering factor along the horizontal axis (the real component of the wave vector in an Argand diagram) and  $if''$  along the vertical axis (the imaginary component). The difference resulting in the structure factors  $F(hkl)$  and  $F(\bar{h}\bar{k}\bar{l})$  can be used to identify the location of the heavy atom in the unit cell. The

difference between the amplitudes of the structure factors of a Bijvoet pair of reflection  $hkl$  and  $\overline{hkl}$  is as follows:

$$\Delta|F|_{ano} = \left[ |F_{PH}(hkl)| - |F_{PH}(\overline{hkl})| \right] \frac{f'}{2f''}$$

Analogous to the procedure in the MIR technique, the location of heavy atoms can be derived by calculating anomalous Patterson maps with  $(\Delta|F|_{ano})^2$ .

#### 3.6.8.5 Solution of the phase problem for RcpB by MAD

The MAD data sets were collected as described in 3.6.5 and comprised three anomalous data sets: edge, peak, remote. By treating MAD as a special case of MIR, the selenium sites in the SeMet-RcpB crystals were identified from anomalous difference Patterson and Bijvoet difference Fourier maps calculated with the CCP4 suite [Collaborative Computational Project, 1994]. After phasing with MLPHARE and density modification by solvent-flattening with DM, the resulting electron density was of sufficient quality to automatically build a preliminary model of the protein with the program wARP [Perrakis, et al., 1999] as described in 3.6.9.

To solve the structure of native RcpB, the phases from this initial SeMet-derivative structure were combined with the native-RcpB structure factor amplitudes and an initial electron density map was calculated. [3.6.9.1].

#### 3.6.9 Model building and refinement

If the protein structure was solved *ab initio*, like in the case of RcpB, model building started out from an interpretable electron density map. If the determined phases and the resulting map are of reasonable quality and extend to a high resolution (2.2 Å or better), an approximate model can be automatically built by tracing the backbone chain of the protein in the electron density. This was done for SeMet-RcpB (resolution 1.6 Å) with the software ARP/wARP [Perrakis, et al., 1999]. In all other cases (RcpA and Mg-RcpB) where the structure was solved by molecular replacement, the (modified) search model represented the initial model for building and refinement and was modified to match the amino acid sequence of the actual protein if necessary.

Interactive model building was done with the molecular modeling software O [Jones, 1991] on Silicon Graphics Workstations. During this process, amino acid residues in the proteins were manually added, mutated and/or fitted into the electron density and the

overall chemical correctness of the model was evaluated and approved. Some of the water molecules, all ions and the phosphate group were manually positioned.

The structure factors calculated on the basis of an initial model are generally in rather poor agreement with the experimental structure factors. The degree of correlation between calculated and observed structure factors is represented by an  $R$ -factor:

$$R = \frac{\sum_{hkl} \|F_{obs} - k|F_{calc}\|}{\sum_{hkl} |F_{obs}|} \times 100\%$$

Refinement is the process of adjusting the model to better match the observed structure factors and at the same time to preserve stereochemical and chemical correctness. The improvement of the model is monitored primarily by the  $R$ -factor. In principle, refinement involves correction of positional parameters and temperature factors for all atoms, as well as building of water molecules, ions and other ligands or prosthetic groups. The automated refinement of the structures of RcpA and RcpB was done using the crystallographic software suite CNS (Brünger, et al., 1998). To prevent the molecular structures from being over-refined, an unbiased  $R_{free}$ -factor was calculated during the refinement process, based on a random selection of approximately 5-10% unique reflections as a test set  $T$  (Brünger, 1992, Brünger, 1993).

$$R_{free} = \frac{\sum_{hkl \in T} \|F_{obs} - k|F_{calc}\|}{\sum_{hkl \in T} |F_{obs}|}$$

Generally, the refinement and building of molecular structures started with a rigid body refinement to optimize the position of the molecules in the unit cell. At this stage, the molecular structure is regarded as a rigid entity, the temperature factors are set to 15 Å<sup>2</sup>, and bond angles and lengths remain unaffected. The next step included the atomic positions, dihedral angles and temperature factors in the refinement and was carried out by iterative cycles (100-200 cycles) of automated procedures referred to as simulated annealing and least-squares-minimization (Brünger, et al., 1997). Each automated refinement step was followed by temperature refinement (CNS) and interactive building in O, in order to validate and correct the model. In the case of poorly defined electron densities, especially in regions of the model with high temperature factors like solvent exposed loops, parts of the model needed 'from scratch' building into the given electron density. Water molecules were located by investigating positive peaks in  $\sigma_A$ -weighted difference maps  $F_{obs} - F_{calc}$  (Kleywegt and Brunger, 1996, Kleywegt and Read, 1997), with CNS and manually. All automatically built water molecules were

evaluated with respect to distance parameters and chemical plausability and, if need be, were removed from the model.

### 3.6.9.1 Refinement of the structures of RcpA and RcpB

#### SeMet- and native RcpB

After a partial SeMet-RcpB model was automatically built into the electron density map that resulted from MAD phasing [3.6.8.5] with wARP [Perrakis, et al., 1999], manual model building in O [Jones, 1991] and initial refinement in CNS [Brünger, et al., 1998] gave a preliminary  $R$ -factor of 30.0% and an  $R_{free}$  of 31.4%.

The native structure of SeMet-RcpB was completed by iterative cycles of rebuilding and refinement against data from 20.0-1.75 Å to a final  $R$ - and  $R_{free}$ -factor of 20.0% and 24.1%, respectively. Water molecules were built automatically in CNS as well as manually. Except for loops H2-β3 and loop β5-H5 which needed extensive manual building, the interpretation of the electron density maps was straightforward. The final model comprises all residues 1 to 149 in molecules A and B, with the initiating methionine replaced by alanine.

#### RcpA

Starting with the initial model from the MR procedure,  $2F_{obs}-F_{calc}$  and  $F_{obs}-F_{calc}$  maps were calculated and successive cycles of model building and refinement were performed with 5% of the reflections set aside for the calculation of the free  $R$ -factor [Brünger, 1992, Brünger, 1993]. In order to minimize model bias, composite omit maps were calculated and used for critical inspection of the improving model. The calculated  $2F_{obs}-F_{calc}$  and  $F_{obs}-F_{calc}$  maps clearly indicated the presence of additional atoms at residue Asp70 corresponding to a covalently attached phosphoryl group. It was built as a phosphorylated aspartate residue using the respective files from the HIC-Up database [Kleywegt and Jones, 1998]. Automated water building, further cycles of rebuilding and positional refinement to convergence, resulted in an  $R$ -factor of 23.3 and an  $R_{free}$  of 27.7. Except for the terminal residues 1-9 and 149 and the solvent exposed loops H1-β2 and H2-β3, all residues were clearly visible in the electron density. The final RcpA model covers the residues 10 to 148.

### 3.6.10 Validation of molecular structures

Statistical analysis of a molecular model including root mean square deviations (RMSD) of the bond lengths and angles, dihedrals and temperature factors are implemented in the CNS refinement scripts and were calculated at the end of any refinement step. Comprehensive geometrical analysis was additionally performed with PROCHECK V. 2.0 [Laskowski *et al.*, 1993].

### 3.6.11 Visual representation of molecular structures

Interactive model building was done within O [Jones *et al.*, 1991]. Colored and rendered three dimensional illustrations of protein atomic structures were generated with different molecular visualization packages, according to a given requirement. Programs used were MOLSCRIPT [Kraulis, 1991], Insight II, WebLab Viewer Light [Accelrys Inc., [www.accelrys.com](http://www.accelrys.com)], SPDBViewer [Guex and Peitsch, 1997], Molecular surfaces graphics and electrostatic surface potentials were calculated and rendered in GRASP [Nicholls, et al., 1991] and WebLab Viewer Light [Accelrys].

## 4 Results

### 4.1 X-ray Structural Analysis of RcpB

#### 4.1.1 Expression and purification of wild type RcpB

RcpB was expressed in the *E. coli* strain BL21(DE3) Gold as described in 3.3.5. In brief, the plasmid pET28a+ carrying the coding sequence for the response regulator in a genetic context supplying the protein with an N-terminal thrombin recognition site/His<sub>6</sub> peptide, was transformed into the cells by electroporation (3.2.8). Positive transformants were selected on kanamycin agar plates and grown in TB medium to an OD<sub>600</sub> of 0.8-1.0. The expression was induced at 30°C by adding IPTG to a final concentration of 400 μM to the culture and continued for max. 6 hrs. Cell lysis and the purification of RcpB was done as described under 3.3.7. In order to reduce inhomogeneity in the protein sample, the His<sub>6</sub> tag was cleaved off by incubation with thrombin as described (3.3.13) and the sample was subjected to a two-step purification procedure (3.3.7). The quality of the final protein was verified by SDS-PAGE (fig. 4.1) and ESI-MS analysis. According to the elution behavior in size exclusion chromatography, RcpB had an apparent molecular weight of 35 kD in solution.

Nevertheless, the ESI-MS analysis showed that the samples were composed of two main protein species with molecular masses of approx. 17126 and 17208 amu ( $MW_{\text{theor}}$  for RcpB is 17125 amu, presuming the initial methionine is cleaved off following synthesis). A second mass peak with a difference of +82 amu (PO<sub>3</sub>H 79.9 amu) most probably reflected the presence of a phosphorylated form of RcpB. This was verified through *in vitro* phosphorylation by incubating the protein with acetyl phosphate as outlined in 3.3.14. The ESI-MS analysis after phosphorylation showed that the mass peak originating from the phosphorylated species (17206 amu) became more prominent whereas the mass peak for the apo form (17127 amu) diminished noticeably.

After purification and thrombin cleavage the protein was sufficiently pure from other proteins from the cell extracts, to be subjected to crystallization. One liter of a standard expression culture yielded approx. 10 mg of purified, recombinant protein after thrombin cleavage.

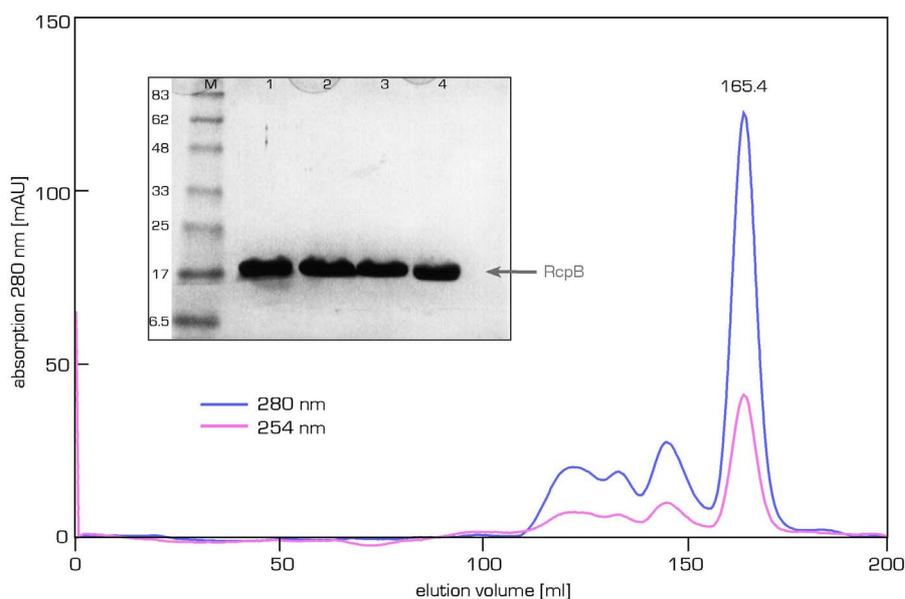


Figure 4.1. **Purification of wild type RcpB.** The chromatogram shows the elution profile of wild type RcpB after thrombin-catalyzed cleavage of the N-terminal His<sub>6</sub> tag from a HiLoad 26/60 Superdex 75 prep grade column. Flow rate 0.53 ml min<sup>-1</sup>, elution buffer G (150 mM NaCl, 5 % glycerol, 50 mM Tris HCl pH 8.0). According to the column calibration (see 3.3.9), the elution volume of the main peak at 165.4 ml corresponds to an apparent molecular weight of approx. 33.4 kD, twice the weight for monomeric RcpB (17.1 kD). The inset shows the coomassie stained SDS-PAGE analysis of the pooled fractions (1-4, same sample). The denatured protein runs at the expected height near the 17 kD marker band (M). Protein samples were mixed with an equal volume of sample buffer (3.3.2), heated to 80°C for 5 min and applied to the gel. The gel was run at constant power (180 V) until the lower dye band of the sample buffer reached the lower end of the gel.

#### 4.1.2 Expression and purification of SeMet-RcpB

For a successful structural characterization employing the MAD technique as described in 3.6.8.4, RcpB had to be labeled with the analog amino acid L-selenomethionine in order to introduce an anomalous scatterer into the protein. This was done by expressing the protein in a methionine auxotroph *E. coli* strain, grown in a minimal medium that contained all amino acids except L-methionine, which was substituted by L-selenomethionine [see 3.3.6] [Budisa, et al., 1995, Hendrickson, et al., 1990]. Under appropriate conditions the cells, void of their ability for methionine biosynthesis, readily incorporated the offered, modified amino acid into the proteins. Major drawbacks in this step were the toxicity of L-selenomethionine in higher concentrations, and the fact that the growing behavior of the cells was slightly altered. Both were mastered by using the expression protocol described in 3.3.6. The protein was purified in two steps as described and special attention was paid to maintaining non oxidizing conditions throughout the procedure. After removal of the affinity tag with thrombin and another round of purification, the protein's quality was analyzed by SDS-PAGE and ESI-MS. Fig.

4.2 demonstrates the high purity of the recombinant protein which was sufficient for crystallization. As deduced from the analytical gelfiltration, the molecular weight of SeMet-RcpB in solution corresponds to approx. 35 kD [dimeric RcpB]. The MW determined by ESI-MS was found to be 17266.8 amu. The MW of the unlabeled protein was 17126.4 amu as determined in 4.1.1 and the mass difference of 140.4 amu perfectly matches the complete substitution of all three methionine residues per molecule by the selenium-analog.

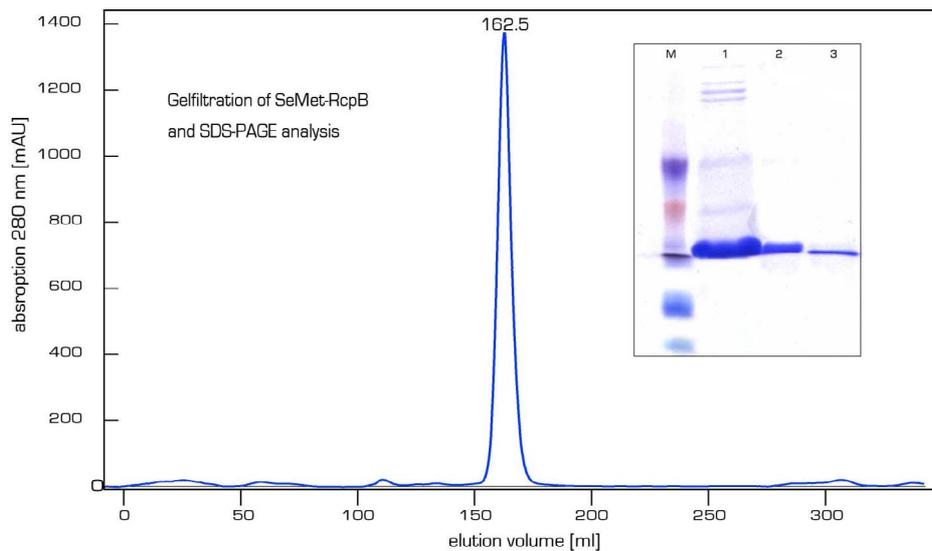


Figure 4.2. **Purification of SeMet-RcpB.** The chromatogram depicts a typical gelfiltration profile (HiLoad 26/60 Superdex 75 prep grade column, flow rate 0.53 ml min<sup>-1</sup>, elution buffer G [150 mM NaCl, 5 % glycerol, 50 mM Tris-HCl pH 8.0]) of SeMet-RcpB after incubation with thrombin and Ni-IMAC purification. Like in the case of wild type RcpB, the elution volume of 162.5 ml corresponds to an apparent molecular weight of approx. 35 kDA indicating dimer formation. The coomassie stained SDS-PAGE analysis demonstrates the purity of the final sample (3) as obtained from the gelfiltration. (M) protein molecular weight marker, lane (1) fractions eluted from Ni-IMAC step, lane (2) sample after gelfiltration and prior to thrombin cleavage.

### 4.1.3 Crystallization and diffraction data collection of RcpB

A preliminary screening for suitable crystallization conditions for RcpB with standard screens (3.6.1) readily resulted in a few possible starting points for optimization. Crystals of the highest quality of both, wild type and SeMet-RcpB, were obtained from precipitating solutions containing 1.75-2.0 M sodium formate (depending on the protein concentration, which was generally in the range of 2-5 mg ml<sup>-1</sup>), 2.5 % glycerol and 0.1 % 2-mercaptoethanol in hanging drops at 4°C (see 3.6.3). Crystals appeared within a few days and reached a final size of approx. 200 microns in each dimension after one week (fig. 4.3). Both, RcpB and SeMet-RcpB formed crystals belonging to the tetragonal space group P422 with the unit cell dimensions  $a = b = 72.2 \text{ \AA}$ ,  $c = 142.8 \text{ \AA}$  and  $\alpha = \beta = \gamma = 90^\circ$ .

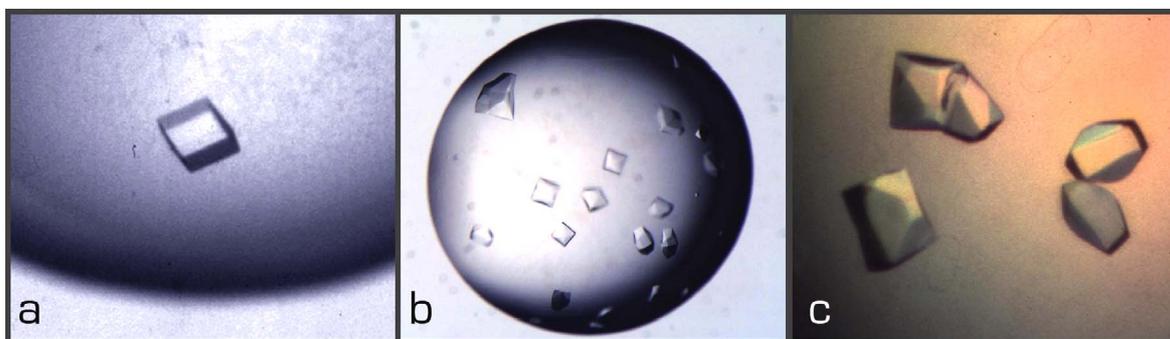


Figure 4.3. **Single crystals of wild type and SeMet-RcpB.** Diffracting single crystals of RcpB were grown in hanging drops under two conditions: (a) 7.5 % PEG, 50 mM MgCl<sub>2</sub> and (b) 1.75-2.0 M sodium formate, 2.5 % (v/v) glycerol and 0.1 % 2-mercaptoethanol. The resulting crystals belonged to spacegroups P2<sub>1</sub> (a) and P4<sub>2</sub>:2 (b) For the structure determination, only the tetragonal crystals were used. (c) SeMet-RcpB crystallized in the same precipitation buffer as the wild type RcpB and formed tetragonal crystals (P4<sub>2</sub>:2) of approx. 100 microns in each dimension.

#### Data collection of wild type RcpB

A native data set from a single cryo-cooled crystal was collected at DESY beamline BW6 (Hamburg) to a resolution of 1.75 Å. The crystal was rotated in two sweeps of contiguous oscillation ranges of 1°, separated by 90°. Each sweep comprised a total range of 30° as determined with the STRATEGY option included in the MOSFLM software package (Leslie, 1992). Figure 4.4 shows a typical diffraction image recorded with a MAR CCD camera.

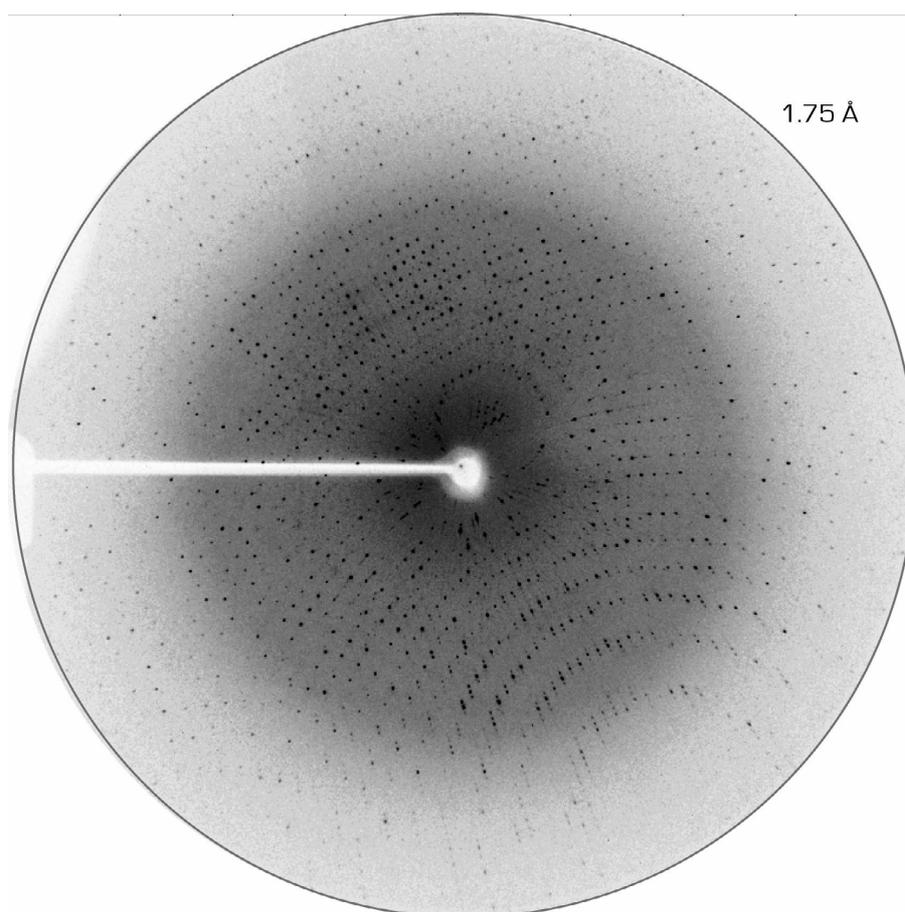


Figure 4.4. **Oscillation image of a tetragonal wild type RcpB crystal** recorded on a MAR CCD detector at the DESY synchrotron beamline BW6. The resolution at the outer rim of the image is 1.75 Å.

Data reduction was performed as described in 3.6.6. The overall completeness of the data was 98.2 % with an  $I/\sigma I$  of 12.2 in the highest resolution shell (1.78-1.75 Å). An overview of the data collection statistics is given in table 4.1. From systematic absences in the reflections, the nature of the four-fold screw axes were determined to be either  $4_1$  or  $4_3$  along  $a$ . This implied that, under the employed conditions, RcpB crystallized in space group  $P4_12_12$  or the enantiomorphous  $P4_32_12$

Data collection and refinement statistics of RcpB	
<b>Crystal</b>	RcpB
Space group	P4 <sub>2</sub> :2 <sub>1</sub> :2
Cell dimensions ( <i>a,b,c</i> [Å])	72.2, 72.2, 142.8
Resolution limit [Å]	64.5-1.75
Wavelength [Å]	1.009
Measured Reflections (total/unique)	260,719/39,024
Completeness (overall; last shell; %)	98.2/98.8
$R_{\text{merge}}^1$ (overall; last shell; %)	3.9/12.2
$I/\sigma$ (overall; last shell; %)	27.4/12.2
$B$ -factor (Wilson/overall mean) <sup>2</sup>	22.2/26.0
Final $R$ -factor/ $R_{\text{free}}$ [%]	19.9/24.3
RMS Deviation from ideal geometry:	
Bond [Å]	0.005
Angles [deg]	1.3
Dihedrals [deg]	22.9

<sup>1</sup>The resolution range in the highest bin (last shell) was 1.78-1.75 Å for RcpB

<sup>2</sup>The  $B$ -factor was estimated from a Wilson plot; for the native datasets over the resolution range 3.9-1.75 Å

Table 4.1

### Data collection and refinement statistics of SeMet-RcpB

For the MAD experiment, a total of four data sets at different wavelengths were collected at the BW6 beamline (Hamburg). For this purpose, a single crystal was mounted to the goniometer head of the diffraction apparatus and flash frozen to 100 K. Prior to data collection, the exact wavelengths for the MAD experiment were determined by wavelength scanning around the theoretical K absorption edge of selenium (K-edge 0.9795 Å), and measuring the X-ray fluorescence scattering from the mounted crystal. From the obtained X-ray fluorescence spectrum, the following wavelengths were chosen for the data collection:

Dataset	Wavelength [Å]
Peak	0.98
Inflection	0.9811
Remote	0.95
Native	1.009

All four data sets (peak, inflection, remote, native) were collected from a single crystal and to a resolution of 1.60 Å. The crystal was rotated in steps of 0.5° through two angular ranges [51 to 111° and -131 to -71°], each comprising 60°. The reflections were recorded with a constant dose of 400 kHz sec<sup>-1</sup> on a MAR CCD camera. Data reduction was done with the software package HKL as described (3.6.6). Table 4.2 gives a summary of the data collection parameters and statistics.

As determined from the processed diffraction data, SeMet-RcpB crystallized in the tetragonal space group P4<sub>2</sub>2 with either a 4<sub>1</sub> or 4<sub>3</sub> four fold screw axes along the *a* cell axis (P4<sub>1</sub>2<sub>1</sub>2 or P4<sub>3</sub>2<sub>1</sub>2) and two molecules in the asymmetric unit (the Matthews coefficient was 2.7 for an estimated solvent content of 54 %, calculated with MATTHEWS\_COEFF as part the CCP4 suite; see 3.##.).

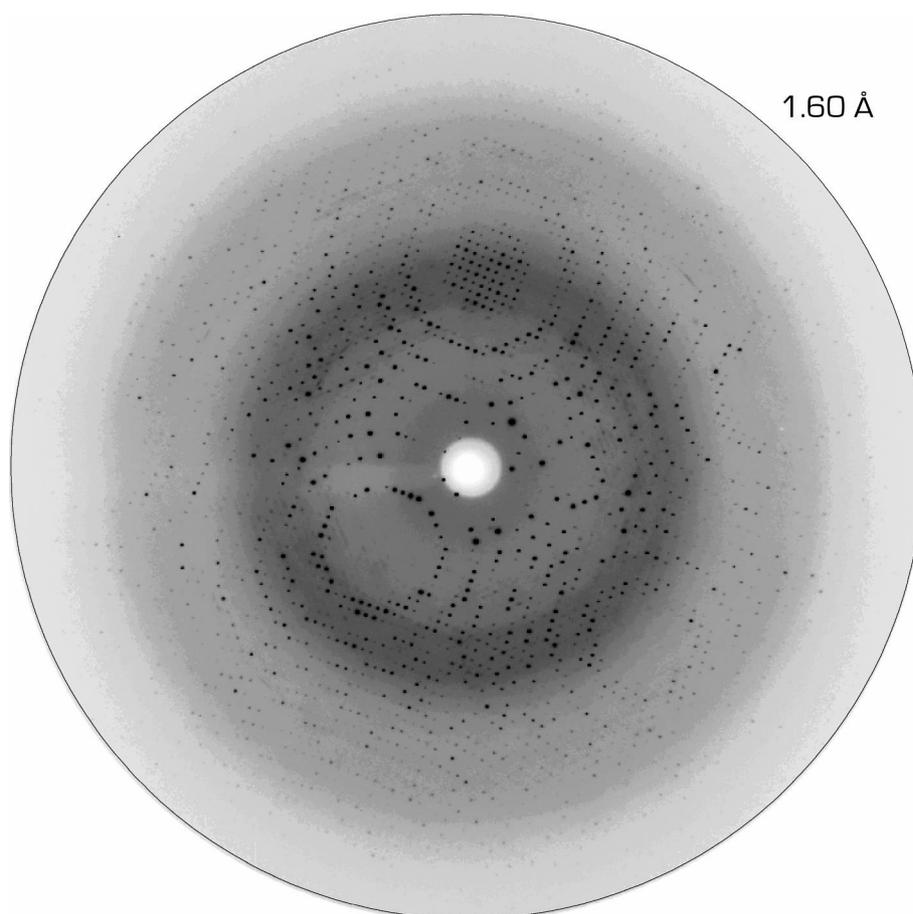


Figure 4.5. **Oscillation image of a tetragonal SeMet-RcpB crystal.** The image was recorded on a MAR CCD detector at the DESY synchrotron beamline with a constant X-ray dose of 400 kHz min<sup>-1</sup>. The resolution at the outer rim of the image is 1.60 Å.

#### 4.1.4 Determination of the phase angles for RcpB by MAD

The phases for the SeMet-RcpB data set were determined by the multiple anomalous wavelength dispersion technique as generally explained in 3.6.8.4. The software routines used were part of the CCP4 package. The anomalous differences at the peak wavelength were used to calculate anomalous difference Patterson maps with FFT (CCP4). From these maps, heavy atom positions were determined by peak search in Harker sections (PEAK\_SEARCH, CCP4) and subsequent refinement of the positions with MLPHARE. The problem of the enantiomorph ambiguity in the location of the heavy

atom sites was solved by simple inspection of resulting initial electron density maps. Only the correct handed solution gave an interpretable map. The phase calculation for all structure factors were performed with MLPHARE followed by density modification [solvent flattening] with DM [CCP4].

#### 4.1.5 Model building and structure refinement of RcpB

Due to the high resolution and quality of the calculated electron density an initial model could be built using the automated routine ARP/wARP [Perrakis, et al., 1999], which traces and refines the backbone atoms of a given protein in the electron density map, and thereby simplifies the otherwise tedious task of manual building from scratch. The model derived from the wARP routine was still incomplete [276 residues of 294 of the homodimeric molecule] and consisted of 5 chains where two were expected [connectivity index 0.96]. In the following process of manual model building using O [Jones et al., 1992] and automated refinement with CNS, the model was preliminarily improved to such an extent [ $R$ -factor of 30.0 % and an  $R_{free}$  of 31.4 %.] that the phase information was suitable to be combined with the structure factors of the wildtype data set of RcbB derived from 4.1.3. In the further process, only the wild type structure was refined and completed and the SeMet-RcpB structure was only consulted in cases where the positioning of disordered groups/atoms was ambiguous. In iterative cycles of interactive model building, positional refinement by simulated annealing and energy minimization with CNS against data from 20 to 1.75 Å, and automated and manual water building, the molecular model could be completed with a final  $R$ -factor and  $R_{free}$  of 20.0 % and 24.1 %, respectively. Except for loops H2/ $\beta$ 3 and loop  $\beta$ 5/H5, which needed extensive manual building due to their solvent exposed location in the molecule which resulted in a poorly defined density and high  $B$ -factors, the interpretation of the electron density maps was straightforward. The final model comprised all residues 1 to 149 in both monomeric subunits, with the initiating methionine replaced by alanine. A total of 474 water molecules were built and one 2-mercaptoethanol C<sub>2</sub>H<sub>6</sub>OS, that formed a disulfide bridge with cysteine side chain C55. The crystal structure of RcpB was deposited in the Protein Data Bank [[www.rcsb.org](http://www.rcsb.org)] under the PDB ID 1K66.

Data collection and refinement statistics for RcpB and SeMet-RcpB										
<b>Crystal</b>	RcpB		SeMet-RcpB							
Space group	P4:2:2									
Cell dimensions (a,b,c (Å), $\alpha$ , $\beta$ , $\gamma$ )	72.2, 72.2, 142.8, 90, 90, 90									
<b>Data collection (cryogenic)</b>	Native	Edge	Peak	Remote						
Resolution limit (Å)	64.5-1.75	19.8-1.60	19.8-1.60	19.2-1.60						
Wavelength (Å)	1.009	0.9811	0.9800	0.9500						
Measured Reflections	260,719	1,088,308	1,092,331	1,092,368						
Unique reflections	39,024	50,575	50,823	50,756						
Completeness <sup>1</sup> (overall; highest bin %)	98.2	99.1/99.6	99.0/97.5	99.1/99.7						
R <sub>merge</sub> <sup>1</sup> (overall; highest bin; %)	3.9/12.2	3.5/30.0	3.7/29.4	3.6/31.7						
I/ $\sigma$ I (overall; highest bin; %)	27.4	42.7/5.6	38.8/5.6	39.6/5.4						
B-factor (Wilson) <sup>2</sup>	22.9	21.4	21.0	21.1						
<b>Phasing (MLPHARE)</b>										
Resolution shell (Å)	8.18	5.15	3.76	2.96	2.44	2.08	1.81	1.60		
Mean Figure of Merit	0.5439	0.6170	0.5512	0.5318	0.5248	0.4345	0.2909	0.1488		
Total Figure of Merit	0.3448									
<b>Refinement (RcpB)</b>										
Number of molecules in AU	2		Number of reflections (work/test)				37955/3771			
Average B-factor (Å <sup>2</sup> )	26.05		Test set selection				shells			
Number of amino acids	298		Final R-factor/R <sub>free</sub>				21.6/25.0			
Number of solvent molecules	393		R.m.s. deviation from ideal geometry							
Resolution used (Å)	15.0-1.75		bond (Å)				0.005			
			angles (°)				1.3			
			dihedrals (°)				22.9			

<sup>1</sup>The resolution range in the highest bin was 1.78-1.75 Å for the native data set and 1.63-1.60 Å for the MAD data sets.

<sup>2</sup>The B-factor was estimated from a Wilson plot; for the native data set over the resolution range 3.9-1.75; for the MAD data sets over the range 3.9-1.6 Å.

**Table 4.2**

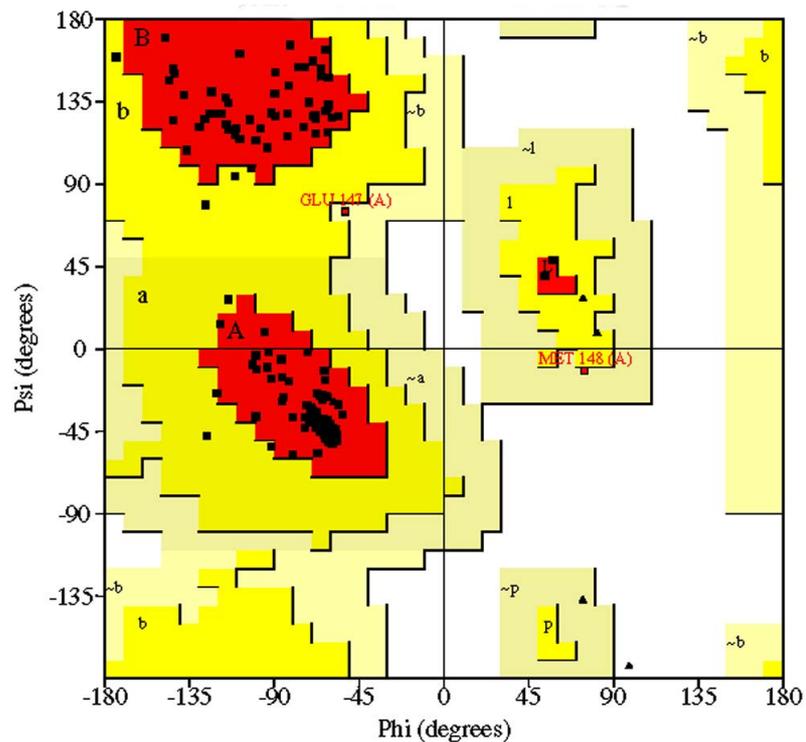


Figure 4.6. **Ramachandran diagram** for the refined model of the cyanobacterial RR RcpB (generated with PROCHECK). Most favored regions are colored red, additionally allowed regions yellow and disallowed regions white. 92.4 % of the residues are in most favored conformations.

Residues in most favored regions [A,B,L]	121	92.4 %
Residues in additionally allowed regions [a,b,l,p]	8	6.1 %
Residues in generously allowed regions [~a, ~b, ~l, ~p]	2	1.5 %
Residues in disallowed regions	0	0.0 %
Number of non-glycine and non-proline residues	131	100.0 %
Number of end-residues (excl. GLY and PRO)	1	
Number of glycine residues (shown as triangles)	7	
Number of proline residues	10	
Total number of residues	144	

The crystal structure of RcpB was deposited in the Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)) under the PDB ID 1K66.

## 4.2 X-ray Structural Analysis of RcpA

### 4.2.1 Expression and purification of recombinant RcpA

The heterologous expression of RcpA in *E. coli* cells was accomplished by transforming the strain BL21(DE3) Gold with the recombinant Plasmid pMEX8, hosting the *RCPA* gene. The vector supplied the recombinant protein with a C-terminal His<sub>6</sub> peptide for facile metal affinity purification. High levels of expression of up to 50 mg l<sup>-1</sup> cell culture were obtained at 30°C within 4-6 hrs after induction with 400 μM IPTG. The two step purification (3.3.7) yielded a protein of excellent purity as judged from SDS page (Fig. 4.7). The ESI-MS analysis certified an average molecular masses of 17635 amu and 17714 amu (M+PO3<sup>3-</sup>+H) (M<sub>theor</sub> = 18264.84 kD). According to the elution behavior of RcpA on a Superdex 75 gel filtration column the apparent molecular weight in solution was approx. 35 kD, which corresponded to the dimeric protein's molecular weight (Fig. 4.7).

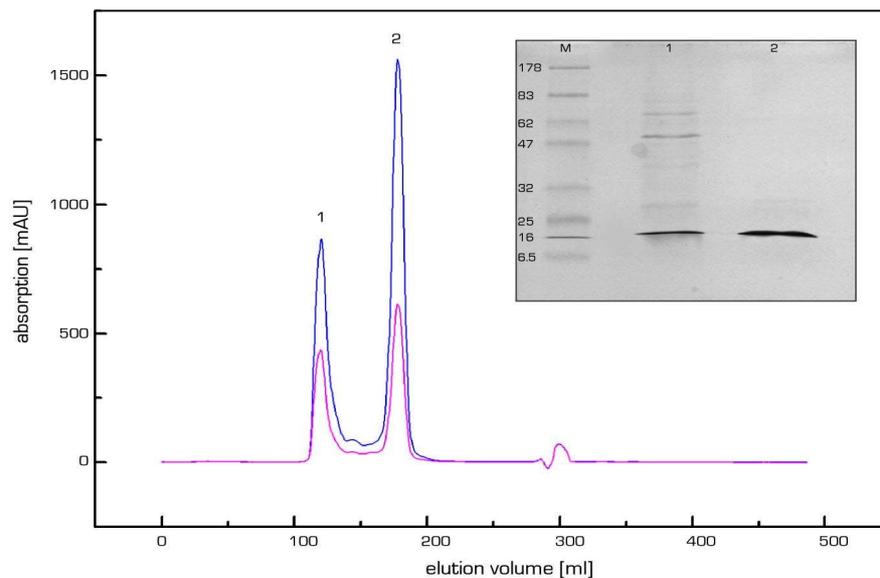


Figure 4-7. **Purification of heterologously expressed RcpA.** Depicted is the chromatogram of a typical size exclusion chromatography on a HiLoad 26/60 Superdex 75 prep grade column, flow rate 0.53 ml min<sup>-1</sup>, elution buffer G (150 mM NaCl, 5 % glycerol, 50 mM Tris-HCl pH 8.0). The sample was injected at 0 ml. Peak (1) corresponds to the cut-off size > 70 kD, (2) RcpA elutes at 178 ml which correlates with an apparent molecular weight of approx. 33 kD. The coomassie stained SDS PAGE (inset) demonstrates the purity of the protein prior (2) and after (3) gel filtration. (M) standard molecular weight marker. Protein samples were mixed with an equal volume of sample buffer (3.3.2), heated to 80°C for 5 min and applied to the gel. The gel was run at constant power (180 V) until the lower dye band of the sample buffer reached the lower end of the gel.

#### 4.2.2 Crystallization and diffraction data collection of RcpA

Crystals of RcpA were obtained as described in 3.6.1 and 3.6.2 employing the sitting drop vapor diffusion technique. Several conditions yielded in microcrystals, mostly needles and plates (Fig. 4.8.b), which were either too tiny to be measured or had developed properly in only one or two dimensions. One three-dimensional single crystal of sufficient size was obtained at 18°C using a well solution that contained 1.4 M sodium citrate, 150 mM sodium chloride, 10 % glycerol and 100 mM HEPES (pH 7.0). The crystallization droplets contained the RcpA solution described in 3.3.7.1, mixed with an equal volume of the well solution. The crystal appeared after two weeks and grew to a size of approx. 100x100x30 microns (Fig. 4.8.a).

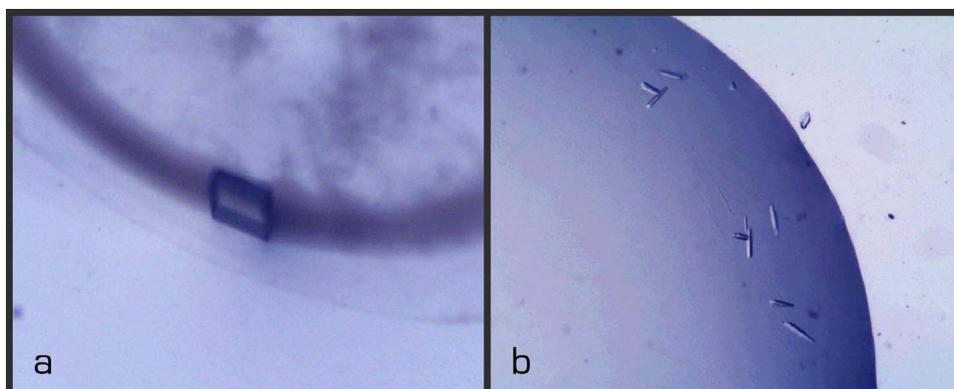


Figure 4.8. **Crystals of the response regulator RcpA.** (a) One orthorhombic crystal was obtained from 0.7 M sodium citrate as described. Diffraction was to a resolution of 1.9 Å. (b) Typically, RcpA crystallized as small needles that were unsuitable for X-ray diffraction.

#### Data collection

For the collection of a diffraction data set, the crystal was mounted on a cryo loop and flash frozen in a cold nitrogen stream as explained in 3.6.5. A complete data set to a resolution of 1.9 Å was collected on a 13.5 cm Ø MAR CCD detector using synchrotron radiation at the BW6 beamline (DESY Hamburg). For data collection, the crystal was rotated through an angle of 0.5° per image around an axis perpendicular to the X-ray beam. The data set comprised 160 images, covering an angular range of 80° composed of two sweeps (0-45°, 50-85°; crystal orientation). Figure 4.9 shows a typical diffraction image.

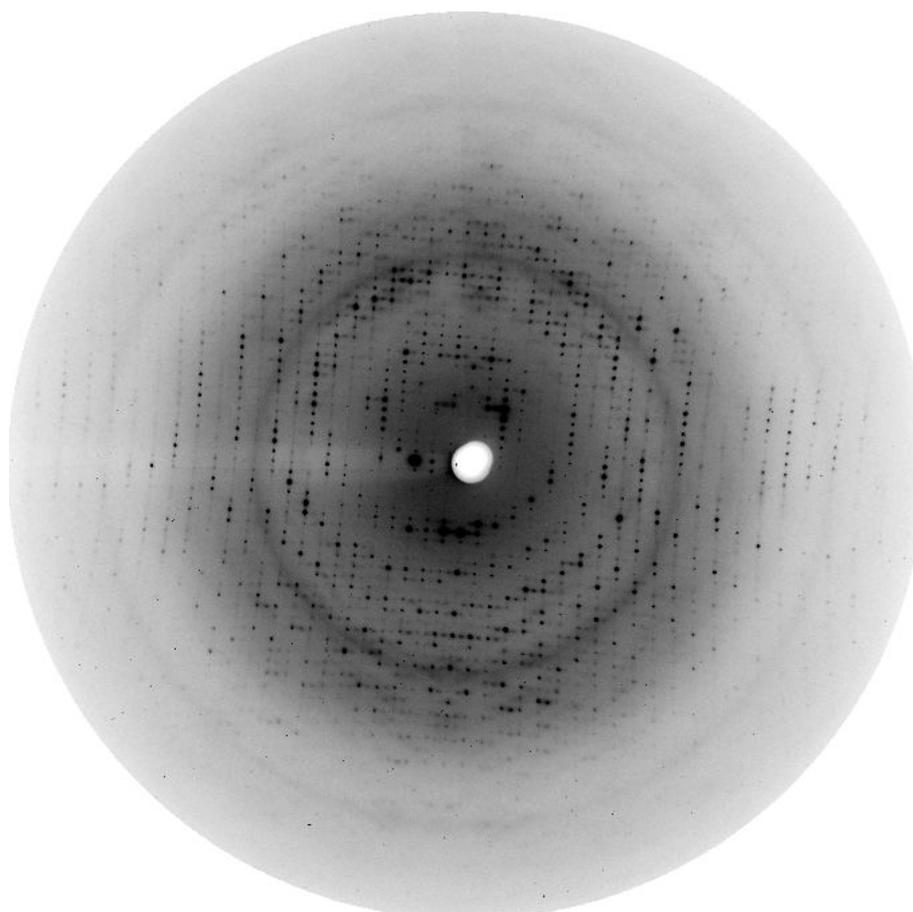


Figure 4.9. **Oscillation image of the orthorhombic RcpA crystal** recorded on a MAR CCD detector at the DESY synchrotron beamline. The resolution at the outer rim of the image is 1.9 Å.

As followed from the crystallographic analysis with DENZO, RcpA crystallized in the orthorhombic space group  $P2_12_12$  with the unit cell dimensions  $a = 76.7$  Å,  $b = 85.2$  Å,  $c = 44.1$  Å and  $\alpha = \beta = \gamma = 90^\circ$ . Data reduction was accomplished with DENZO and SCALEPACK. The two-fold screw axes along  $a$  and  $b$  were deduced from systematic absences in the reflection data. The completeness of the data set was 96.7 % (highest resolution bin 94.6 %) with an  $R_{\text{merge}}$  of 0.037 (0.302) for all HKL. For a summary of the data collection statistics, refer to table 4.3.

#### Determination of the number of molecules in the asymmetric unit (AU)

The number of molecules per AU was determined unambiguously with MATTHEWS\_COEFF (CCP4; see 3.6.#). With an estimated solvent content of approx. 42 %, the Matthews coefficient  $V_M$  was calculated to be  $2.1 \text{ \AA}^3/D$  for two molecules in the AU, which perfectly matches the range found by Matthews (Matthews, 1968). For one or three molecules in the AU, the solvent content was calculated to be 70.7 % and 12.1, and  $V_M$  to be 4.1 and 1.4, respectively, with neither being reasonable for protein crystals.

### 4.2.3 Solution of the phase problem for RcpA by Patterson search

The initial phases for the molecular structure of RcpA were approximated by the Patterson Search or Molecular Replacement (MR) method as outlined in 3.6. The model structure used was a modified model structure of RcpB which was solved as described in section 4.1. All residues in RcpB had been replaced by serines and some by alanines at positions where an alanine was expected in the sequence, RcpA shares approx. 60 % homology with RcpB but only 24 % with the structurally closest homologue, CheY from *E. coli*. The latter proved unsuitable as a search model, even in various modified forms. The rotational and translational searches were performed with the program AmoRe (Navazza, 1994) in the resolution range 20.0–3.0 Å. A best solution was found for two molecules with a correlation coefficient  $CC-F$  of 53.5 % and an  $R-F$ -factor of 45.4 % [ $CC-F$  and  $R-F$ -factor of the next best solution were 45.4 and 51.4, respectively].

### 4.2.4 Structure refinement of RcpA

From the initial model derived from the molecular replacement routine,  $2F_{obs}-F_{calc}$  and  $F_{obs}-F_{calc}$  maps were calculated followed by successive cycles of manual model building and automated refinement as described in 3.6.9. For the calculation of an unbiased free  $R$ -factor, 5 % of the reflections were set aside as the test set during the refinement. In order to minimize model bias, composite omit maps were calculated and used for the critical inspection and correction of the model.

The calculated  $2F_{obs}-F_{calc}$  and  $F_{obs}-F_{calc}$  maps clearly indicated the presence of additional atoms at residue D70 corresponding to a covalently bound phosphoryl group. According to the protein's function and the fact, that this particular aspartate is identical to the highly conserved active site aspartate in other RRs that act as the phosphoryl acceptor site, it was self-evident to assume a covalently bound phosphate. As a consequence, the aspartic amino acid at this position was exchanged for a phosphorylated aspartate residue (PHD) using the respective entries from the HIC-Up database (Kleywegt and Jones, 1998) (Fig. 4.10). The group fitted perfectly into the density as depicted.

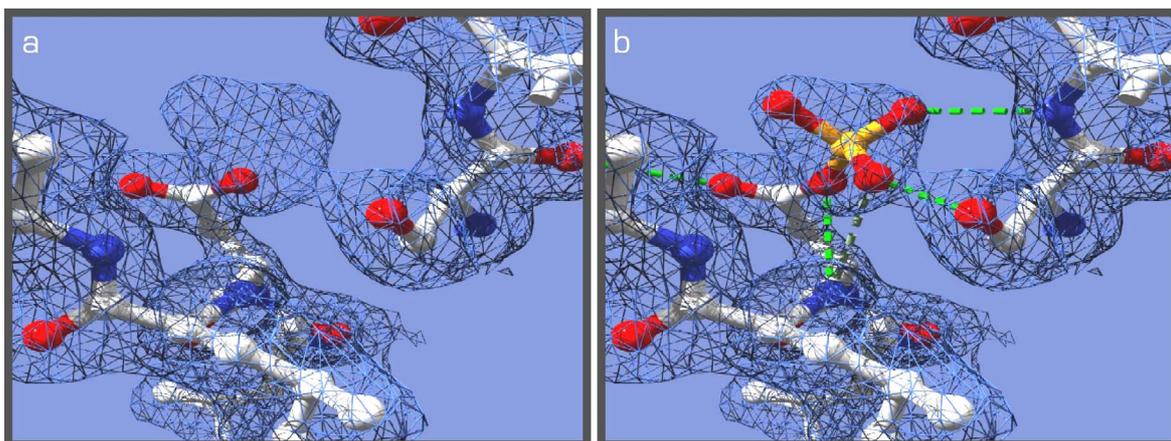


Figure 4.10. **Rendered representation of the active site aspartate D70** with the enclosing electron density contoured at  $1 \sigma$ . Atoms are colored according to type: C (white), O (red), N (blue), P (yellow). The extra density fused to the carboxyl group of D70 (a) perfectly accommodates the covalently bound phosphoryl group (b). The established hydrogen bond network is illustrated as green dashed lines.

The putative metal binding site in the conserved active site was found to be occupied by an octahedrally coordinated metal ion and several water molecules. As deduced from the electrolyte composition of the crystallization condition and especially from the geometry of the coordination sphere of this ion, as well as from comparison with other X-ray structures, it became evident that the atom was a sodium ion. The ion is coordinated by two water molecules, the backbone carbonyl oxygen of N72 and the carboxylate oxygens of E13, D14 and PHD70 in a distorted octahedral fashion (Fig. 4.19). The distances between ligand atoms and the central ion are between 2.27 and 2.73 Å. The distances for equally coordinated magnesium ions generally fall in the range of 2.01 and 2.20 Å whereas those for sodium ions are longer and between 2.10 and 2.62 Å, as found in pdb entries from the HIC-Up data base.

Automated water building, further cycles of rebuilding and positional refinement to convergence resulted in an  $R$ -factor of 23.3 % and an  $R_{free}$  of 27.7 %. Except for the terminal residues 1-9 and 149 and the solvent-exposed loop H1/ $\beta$ 2 and loop H2/ $\beta$ 3, all residues were clearly visible in the electron density. The final RcpA model covered the residues 10 to 148 and included the covalently bound phosphoryl group, the active site sodium ion and 115 water molecules. Residues 1-9 and 149 as well as the N-terminal His6 tag were not visible in the electron density maps, presumably due to disorder or, in the case of the N-terminal peptide, unspecific proteolysis. A few other side chains were not defined in the electron density map and were replaced by glycine or alanine as listed:

K31 (A), A59 (G), R79 (A), I103 (A), E105 (A), R124 (A), E138 (A), S148 (A).

Data collection and refinement statistics of RcpA	
<b>Crystal</b>	RcpA
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2
Cell dimensions [ <i>a,b,c</i> [Å]]	44.1, 76.7, 85.2
Resolution limit [Å]	25.0-1.90
Wavelength [Å]	1.050
Measured Reflections (total/unique)	301367/22731
Completeness (overall/ last shell; %)	96.9/94.6
$R_{\text{merge}}^1$ (overall; last shell; %)	3.7/30.2
<i>I</i> / $\sigma$ (overall; last shell; %)	25.2/3.1
<i>B</i> -factor (Wilson/overall mean) <sup>2</sup>	26.4/38.2
Final <i>R</i> -factor/ $R_{\text{free}}$ (%)	23.3/27.7
RMS Deviation from ideal geometry:	
Bond [Å]	0.005
Angles [deg]	1.3
Dihedrals [deg]	22.7

<sup>1</sup>The resolution range in the highest bin was 1.93-1.90 Å for RcpA.

<sup>2</sup>The *B*-factor was estimated from a Wilson plot; for the native dataset over the resolution range 3.22-1.90 Å;

Table 4.3

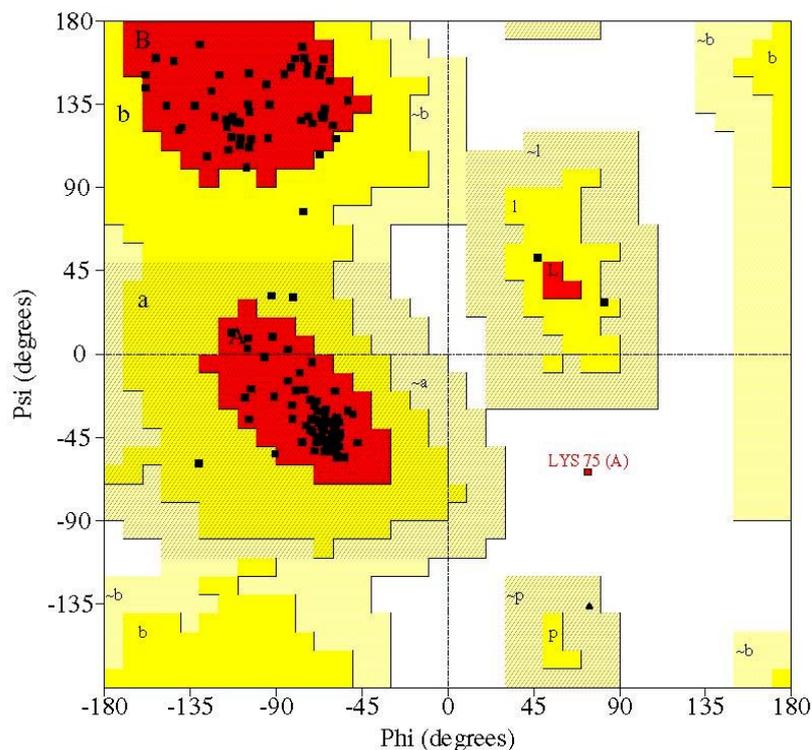


Figure 4.11. **Ramachandran diagram** for the refined model of the cyanobacterial RR RcpA (generated with PROCHECK). Most favored regions are colored red, additionally allowed regions yellow and disallowed regions white. One residue, K75, is in a conformation normally disallowed. 92.1 % of the residues are in most favored conformations.

Residues in most favored regions [A,B,L]	117	92.1 %
Residues in additionally allowed regions [a,b,l,p]	9	7.1 %
Residues in generously allowed regions [~a, ~b, ~l, ~p]	0	0.0 %
Residues in disallowed regions	1	0.8 %
Number of non-glycine and non-proline residues	127	100.0 %
Number of end-residues (excl. GLY and PRO)	1	
Number of glycine residues (shown as triangles)	6	
Number of proline residues	6	
Total number of residues	144	

The crystal structure of phospho-RcpA was deposited in the Protein Data Bank ([www.rcsb.org](http://www.rcsb.org)) under the PDB ID 1K68.

### 4.3 Crystal Structures of RcpA and RcpB

Due to a high overall structural homology of the response regulators RcpA and RcpB, rather than describing each structure separately, a comparative description of the structures of both proteins will be given in the following chapter.

#### 4.3.1 Overall structure of phospho-RcpA and apo-RcpB

Figure 4.12 depicts the overall three dimensional folds of phospho-RcpA. Both RRs crystallized as homodimers in the asymmetric unit, with the respective protomers being interrelated by twofold noncrystallographic rotational axes. The monomeric molecules share the same  $(\beta/\alpha)_5$  architecture with other structurally characterized members of the RR family, i.e. CheY (Stock, et al., 1989) or SpoOF (Lewis, et al., 1999) as predicted from sequence alignments. Basically, they comprise alternating segments of  $\beta$ -strands and  $\alpha$ -helices forming a five-stranded, parallel,  $\beta$ -pleated sheet surrounded by 5  $\alpha$ -helices.

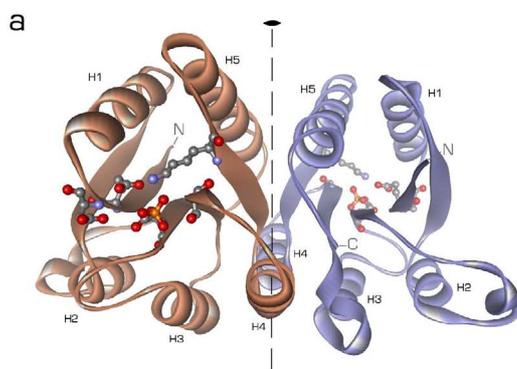


Figure 2a

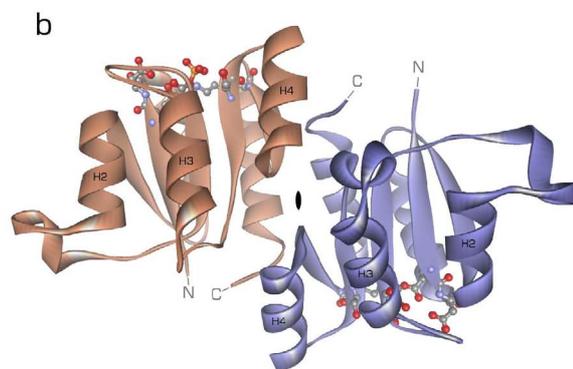


Figure 2b

Figure 4.12. **Overall structure of the RcpA homodimer.** For simplicity, the protein backbones are depicted as ribbons. Dimeric Phospho-RcpA as viewed from (a) top and (b) side with the noncrystallographic twofold rotational axis perpendicular (a) and parallel (b) to the line of vision. The active site in each subunit is drawn as a ball and stick model and is described in detail in 4.3.3. The protein N- and C-termini are labeled. Phospho-RcpA crystallized in space group P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> and diffracted to 1.9 Å. The final model comprised residues 10 to 148; residues 1-9 and 149 and the C-terminal His<sub>6</sub> peptide were not visible in the electron density, presumably due to disorder. The structural information obtained from the SeMet derivative of RcpB was used to solve and refine the structure of apo-RcpA by Patterson search.

The major structural differences between RcpA, RcpB and CheY are shown in figure 4.13, which represents the overlaid backbone traces. The backbone atoms of RcpA and RcpB can be superimposed with a RMSD of 0.99 Å and both can be superimposed on CheY with a RMSD of 1.23 Å [RcpA] and 1.37 Å [RcpB].

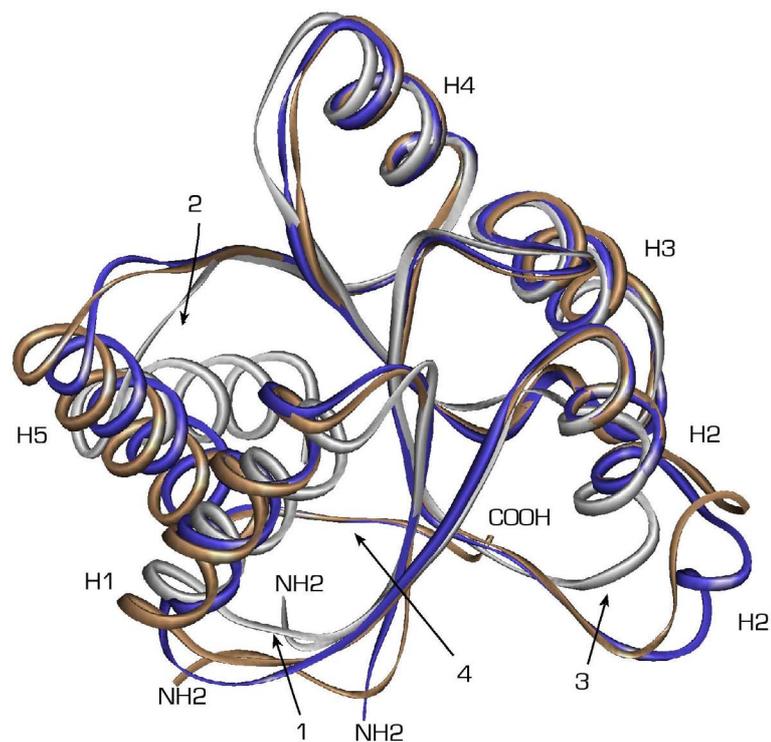


Figure 4.13. **Superimposition of the crystal structures** of RcpA (blue), RcpB (caramel) and CheY (silver). For simplicity, the structures are drawn as backbone worms. The overlay illustrates the overall structural homology and differences in [1] loop  $\beta 1/H1$ , [2] loop  $\beta 5/H5$  and [3] loop  $H2/\beta 3$ . In the case of RcpA, loop  $H2/\beta 3$  consists of residues 52-59 that fold into an additional helical turn ( $H2'$ ). The C-termini of the cyanobacterial RRs [4] are extended with respect to CheY homologues and participate in the formation and stabilization of the homodimers [see paragraph 4.3.2].

As can be seen in the sequence alignment (Fig. 4.14) of the three RRs, the protein sequences of the cyanobacterial RRs have extended regions with respect to the sequence of CheY. In the molecular structure, they result in an extended loop between  $\beta 2$  and  $H3$  [3 in Fig. 4.13] and an extended C-terminus, which is engaged in the formation of specific homodimers as described in the following paragraph.

RcpA	1	MVS	VETE	DKHKK	IFLVED	NKAD	IRLIQE	ALKNS	TVPHEV	V	40
RcpB	1	MVG	- - -	NATQP	LLVVED	SDEDF	SFTFQR	LLQREG	GVVNP	IY	36
CheY	1	-MA	- - -	DKELK	FLVVD	FSTMRR	IVRNLL	LLKELGF	-NNVE		34
RcpA	41	TVRDG	MEAMA	YLRQE	GEYANAS	- - -	RPDL	ILLDLN	LPKKD		77
RcpB	37	RCITG	DQALDF	LYQTGS	YCNPDI	APRPAV	ILLDLN	LPGTD			76
CheY	35	EAEDG	VDA LNK	LQAGGYG	- - - - -	- - - - -	FV	ISDWN	MPNMD		64
RcpA	78	GREVLA	EIKSDPT	LKRIPVV	LSTSI	INEDD	IFHSYD	LHVN			117
RcpB	77	GREVLQ	EIKQDEV	LKKIPVV	IMTTSS	NPKEIE	ICYSYS	IS			116
CheY	65	GLELLK	TIRADG	MSALPVL	MVTAEAK	KENII	IAAAQ	AGAS			104
RcpA	118	CYITK	SRNLS	QLFQ	IVKGI	EEFWL	STATLP	SE -			149
RcpB	117	SYIVK	PLEIDR	LTE	TVQTF	IKYWLD	IVVLP	EMG			149
CheY	105	GYVVK	PFTAAT	LEEKLNK	IFEK	LGM - - -					129

Figure 4.14. **Sequence homology alignment of RcpA, RcpB and CheY.** Two major differences between CheY and the cyanobacterial RRs (indicated by orange bars) which also have a consequence on the three dimensional fold of RcpA and RcpB, are additional residues linking helix 2 and beta sheet 3 and a C-terminal extension,. The insert in RcpA and RcpB forms an additional helical turn H2' at the surface of the protein (see 3 in Fig. 4.13). The C-terminal peptide is involved in dimer formation. The function of the additional helical turn is unknown.

#### 4.3.2 Intermolecular interaction in the homodimers of RcpA and RcpB.

In the crystal structures of RcpA and RcpB, the molecules are associated as distinct dimers in the unit cell with the respective monomers being related by noncrystallographic twofold rotational symmetries (Fig. 4.15). In both RRs, the formation of these homodimeric complexes is mediated by the specific interaction of surface areas spanned by H4,  $\beta$ 5, H5 and the - with respect to CheY - extended C-terminal part of each subunit. As a consequence, the active sites in the dimers reside on opposing positions (Fig. 4.12).

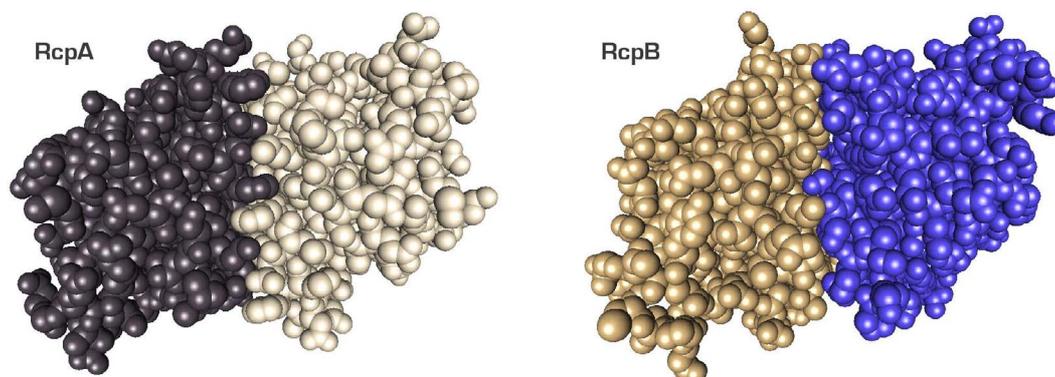


Figure 4.15. **Space filling molecular models** of dimeric RcpA (left) and RcpB (right) as viewed along the twofold noncrystallographic axes, illustrating the overall globular shape of the proteins and the tight and specific interactions between the monomers.

The total interface areas were determined by calculating the differences of Connolly surfaces from two isolated monomers and the dimer (InsightII, probe radius 1.4 Å). They comprise approx. 1000 Å<sup>2</sup> in RcpA and 1175 Å<sup>2</sup> in RcpB or 16 % of the total dimer surface areas. Formation and stabilization of the intermolecular complex is accomplished mainly by the complementary arrangement of a combination of hydrophobic interactions, hydrogen bonds and salt bridges (table 4.4, Fig. 4.16-18). A remarkably large part of this interface is dominated by a central, water-excluding hydrophobic core, that is surrounded by a ring of polar interactions and several bridging water molecules, tying together the rim. The overall design of this interface represents an archetypal protein-protein contact site in homo-oligomeric complexes as described e.g. in a survey by Larsen *et al.* [Larsen, et al., 1998].

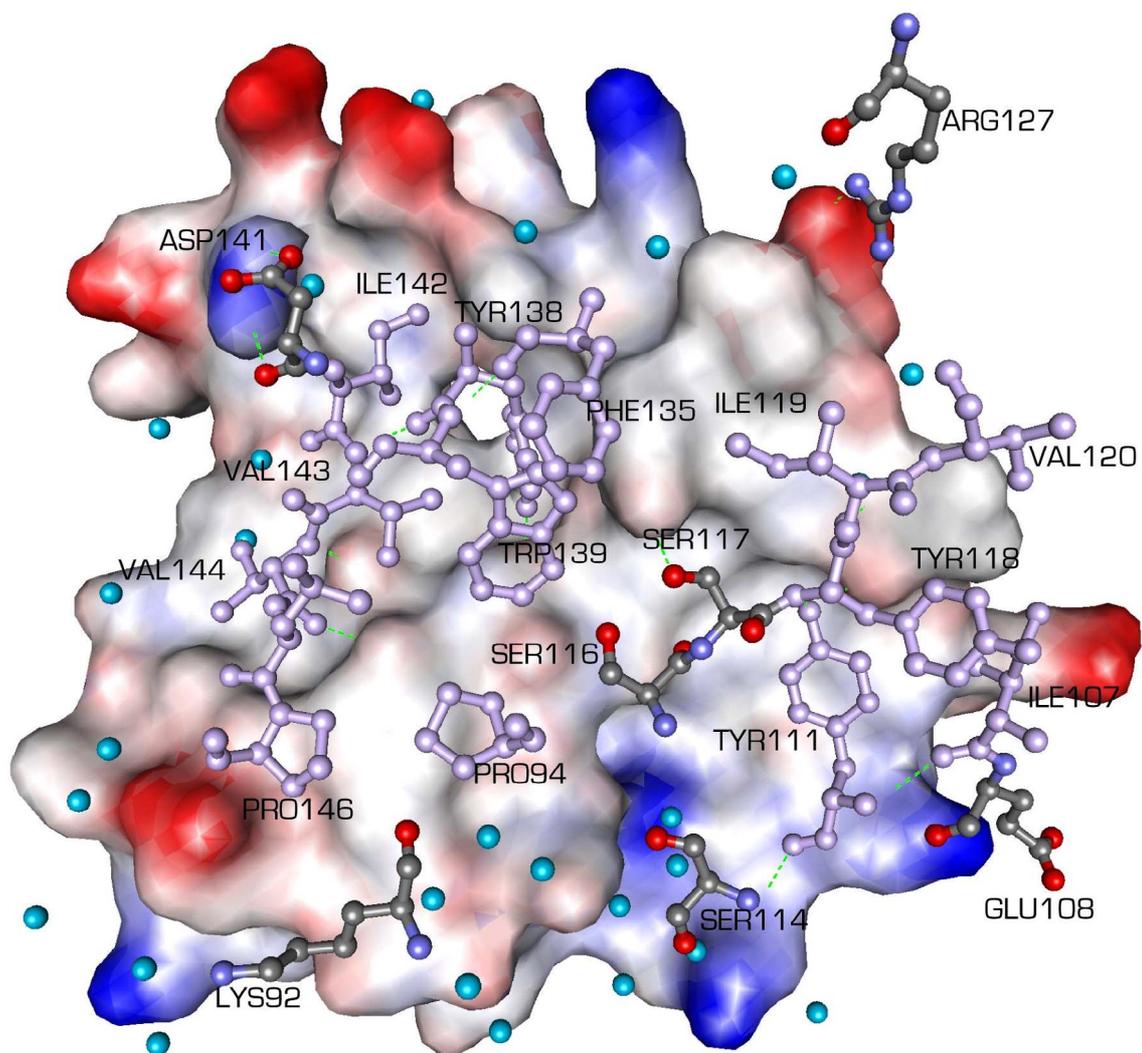


Figure 4.16. **Inter subunit interactions in dimeric RcpA and B.** One monomer is represented by its interacting, solvent accessible surface area, colored according to the electrostatic surface potentials. The interacting side chains of the second monomer are

shown as ball and stick models and colored according to their physicochemical character (non-polar and aromatic: light blue; polar and charged: according to atom type (C (grey), O (red), N (blue)). Interface water molecules participating in the dimer interface formation are depicted as turquoise spheres. A large part of the interface is of mainly hydrophobic nature and excludes water from the surface. The interface rim is populated by water molecules that build up a hydrogen bonded 'seam'.

Within this interface, two prominent, non-polar motifs could be identified that are shared by both, RcpA and RcpB, and that are also conserved in several other RRs (alignment Fig. 4.19), but without exception in RRs of the type associated with OFRs encoding phytochrome-homologous sensor histidine kinases and therefore deserve a more detailed characterization. The most prominent of these conserved motifs, located in the center of the hydrophobic core, is formed by aromatic residues (Y138, W139 in RcpB; F139, W140 in RcpA) in both subunits, that pack against each other to form an arrangement that could be considered an aromatic cluster (4.17). In the case of RcpB, this arrangement is additionally stabilized by a hydrogen bond between the hydroxyl group of tyrosine 138 and the indole NH-group of tryptophan 139. The second motif, located at the rim of the core, comprises two spatially opposing proline residues (P94, P146 in RcpB; P95, P147 in RcpA) from  $\beta$ 4 and the C-terminal loop of one subunit and Y111 (RcpB; Y112 in RcpA) in H4 of the other subunit (Fig. 4.18). The proline pair of one subunit packs against the phenyl side chain of Y111 of the other subunit. The corresponding prolines of the second subunit, in turn, pack against Y111 of the first subunit.

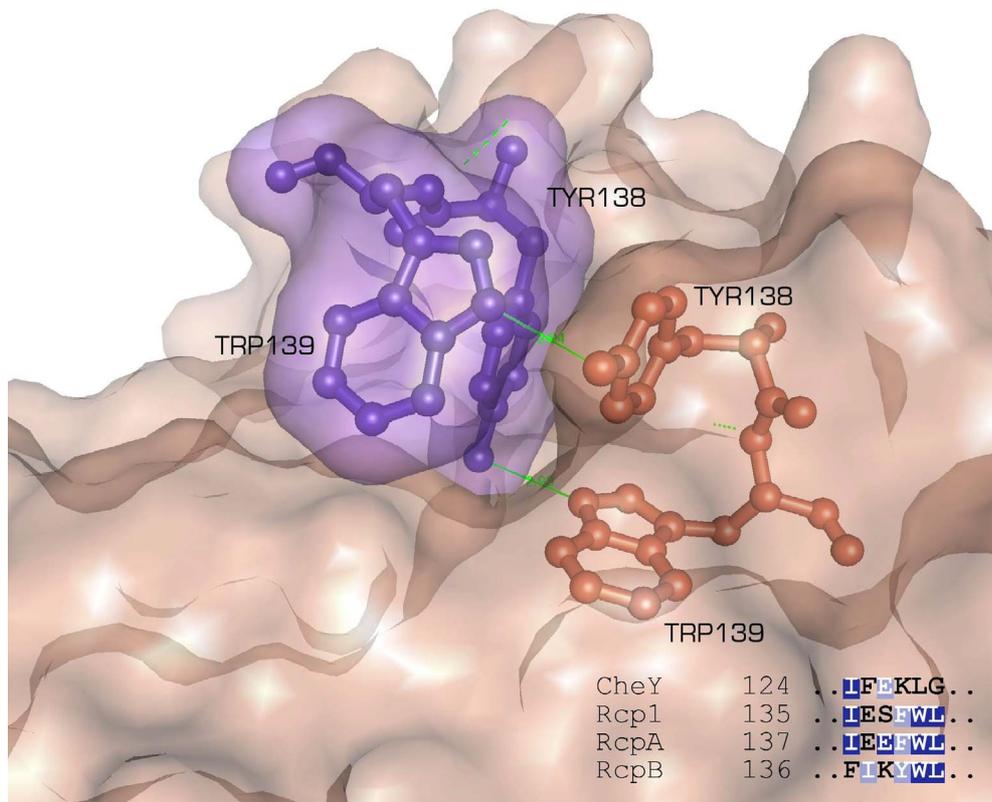


Figure 4.17. **Interface contact in dimeric RcpB.** Close-up view of the complementary packing of the YW motif in dimeric RcpA. The aromatic residues Y138 and W139 of each subunit (orange, blue) are in Van der Waals contact and form a central hydrophobic core in the dimer. The tryptophan hydroxyl group forms a hydrogen bond with the indole NH group (2.93 Å). In RcpA, the equivalent is a phenylalanine/tryptophan pair, as shown in the alignment of conserved interface residues in cyanobacterial RRs (inset).

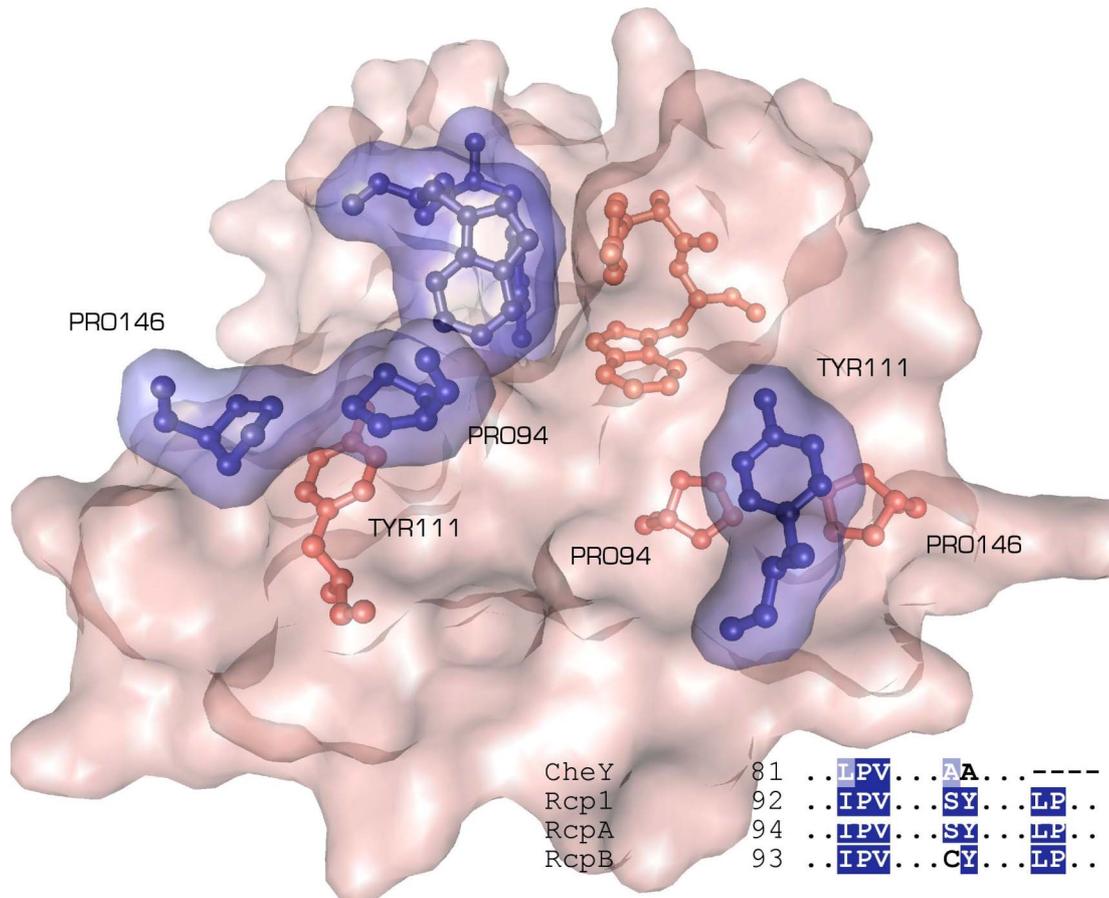


Figure 4.18. **Interface contacts in dimeric RcpA and B.** Close-up view of the complementary packing of the PPY and the YW motifs in dimeric RcpB. One subunit is represented by its solvent accessible surface with the interacting residues shown as ball and stick [orange]. The interacting residues of the second subunit are shown as ball and stick models, enclosed by their Van der Waals surface [blue]. The sequence alignment (inset) shows the conserved residues depicted. RcpA exhibits similar non polar and aromatic interactions.

In RcpB, the dimer hydrophobic interaction additionally comprises two successions of non-polar residues in  $\beta 5$  (Y118, I119, V120, P122) and the C-terminus (I142, V143, V144, L145 [Table 4.4], flanking the aromatic cluster and packing complementarily in the dimer.

In RcpA the hydrophobic character of the C-terminus is not as strongly pronounced as in RcpB and some non-polar interactions are replaced by polar contacts, such as T145, which is H-bridged to Y109 and Y112. Unique to the RcpA dimer interface is the packing of the H4/ $\beta 5$ -loop (subunit A) against H4/ $\beta 5$ -loop (subunit B) which results in the Van der Waals interaction of HIS115 from both subunits.

The remaining non-polar and polar interactions are composed mainly by the complementary arrangement of residues in each subunit in both RRs and are listed in table 4.4. The surface shape complementarity (Lawrence and Colman, 1993) of the interface regions is excellent and in the range of typical protein/protein interfaces,  $S_c = 0.71$  for RcpB and  $S_c = 0.65$  for RcpA [calculated with SC (CCP4); waters not included]. This additionally reflects the strong surface intimacy in the dimers.

RcpA	subunit A	subunit B	RcpB	subunit A	subunit B
<b>polar interaction</b>	LYS92 CE	ASP113 OD2		LYS92 NZ	GLU 108 OG
	ARG93 NE	PHE109 CE1		LYS92 CD	SER112 OG
	ARG93 NH1	ASP113 OD2		GLU108 OE1	ALA147 N
	TYR112 OH	THR145 O		GLU108 OE2	LYS92 NZ
	ASP113 OD1	ARG93 NH1		TYR111 OH	VAL144 O
	ASN117 O	ASN117 ND2		SER114 OG	LYS85 NZ
	ASN117 ND2	ASN117 O		SER116 OG	SER116 O
	TYR119 O	THR144 N		SER117 OG	TYR138 OH
	PHE139 CZ	TRP140 NE1		TYR118 O	VAL144 O
	TRP140 NE1	PHE139 CZ		ARG127 NH2	ASP141 O
	THR145 N	TYR119 O		ARG127 NH1	ASP141 OD1
	THR145 O	TYR112 OH		TYR138 OH	SER117 OG
				TYR138 OH	TRP139 NE
				TRP139 NE	TYR138 OH
			ASP141 OD1	ARG127 NH1	
<b>Hydrophobic interaction</b>	PHE139	VAL96, LE120, ILE136, PHE139		PRO94	TYR111
	PRO95	TYR112		ILE97	TYR138
	PRO147	TYR112		ILE107	VAL144
				TYR111	PRO94, PRO146
	TYR112	PRO95, PRO147		TYR118	VAL144
				ILE119	ILE142, TYR138, VAL143, VAL144
				VAL120	VAL144
				TYR138	TYR138, TRP139
				TRP139	TYR138, TRP139
				ILE142	ILE119
			VAL143		
			VAL144	VAL120, ILE107, TYR118	
			PRO146	TYR111	

**Table 4.4.** List of interface interacting residues in dimeric RcpA and RcpB

### 4.3.3 Comparison with other receiver domains

The sequence information of bacterial phytochromes and their associated phosphate receiver modules were gathered from databases. Together with non-phytochrome associated RRs, they were aligned to see whether interface residues in RcpA and B and Rcp1 are also conserved in other RRs (Fig. 4.19). Strikingly, the alignments revealed a subset of residues involved in the homodimer formation that was highly conserved among different RRs, in all cases from proteins associated with phytochrome-like receptor-histidine kinases. These comprise the cyanobacterial Rcp1 (*Synechocystis*

PCC6803], RcpA and RcpB [*Calothrix* sp. PCC7102] and homologues from *Anabaena* PCC7120 and also RRs from *Deinococcus radiodurans*, and *Agrobacterium tumefaciens*. Additionally, the dimerization motif was found in several other bacterial RRs, some of which are cotranscribed with yet uncharacterized receptor histidine kinases [*D. radiodurans*, *M. thermoautotrophicum*, *A. fulgidus*]. It can therefore be postulated, that phytochrome-associated RRs generally form dimers *via* the described interactions, promoted by a set of invariant [conserved] residues. Therefore, dimerization is a common characteristic of phytochrome-associated RRs, and RcpA, RcpB and homologues constitute a novel subgroup of receiver domains.

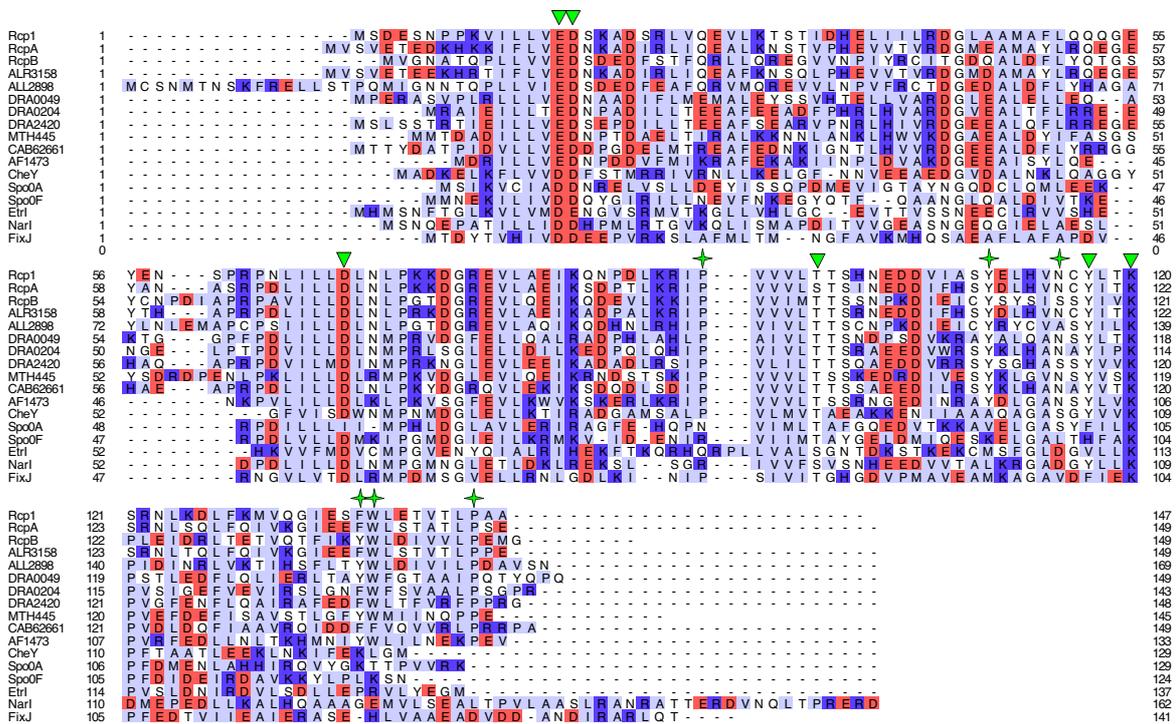


Figure. 4.19. **Homology alignment of RRs of the cheY type.** Highly conserved active site residues are marked with arrow heads. Green stars indicate residues involved in dimer formation in RcpA and B. Rcp1 [*Synechocystis* sp., SLR0474], RcpA and B [*Calothrix* sp.] and RRs from *Anabaena* sp. [ALR3158, ALL2898] and *D. radiodurans* (DRA0049, DRA0204, DR2420) are highly homologous and also co-transcribed with putative bacterial phytochrome-like receptor histidine kinases. Several other RRs, e.g. from *Methanobacterium thermoautotrophicum* [MTH445], *S. coelicolor* [CAB62661] or *Archeoglobus fulgidus*, AF1473] also share the dimerization motifs. On the other hand, none of the structurally characterized receiver domains, like cheY or the homologous SpoDA [*B. stearothersophilus*], SpoOF, Etr1, NarI or FixJ share the motif.

#### 4.3.4 Active site architecture

Like in other cheY-homologous structures, the active site of RcpA and B is an acidic pocket formed by a triad of acidic residues [RcpA: E17, D18, D70, RcpB: E13, D14, D69], a serine [RcpA: S100] or threonine [RcpB: T99] and a lysine [RcpA: K122, RcpB: K121], all together being highly conserved within this family of RR proteins (Fig. 4.20, 4.21; numbering Fig. 4.19).

##### 4.3.4.1 Phospho-RcpA

During the refinement procedure, investigation of  $2F_{obs}-F_{calc}$  and  $F_{obs}-F_{calc}$  electron density maps of RcpA revealed a prominent tetrahedral electron density, fused with the density of the D70 carboxyl group, that could not be accounted for by protein amino acid atoms or solvent molecules (fig 4.10). Due to the biochemical function of this conserved aspartate residue as the phosphoryl accepting group, it was clear and unambiguous that this conspicuous density represented a phosphoryl group, covalently attached to a carboxyl oxygen of D70. The presence of the phosphorylated form of RcpA was already suggested from data derived from ESI-MS spectrometry which reproducibly showed two neighboring masses (17634.8/17714.0 amu). The difference of 79 amu originates from the presence of a covalently attached phosphoryl group. It has been described previously that other RRs were isolated in their phosphorylated form from *E. coli* lysates, either due to unspecific phosphotransfer activity or due to small molecule phosphodonors like acetyl phosphate, the latter being constitutively present in the host cells [McCleary and Stock, 1994].

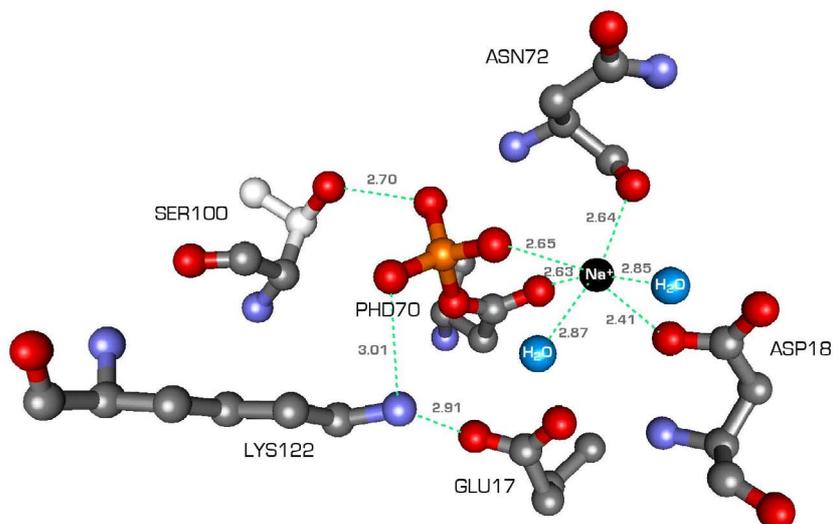


Figure 4.20. **Active site architecture in phospho-RcpA.** Representation of the conserved active site residues, shown as ball and stick models. Atoms are colored according to type: C (grey), O (red), N (pale blue), P (yellow) and water molecules (turquoise). In the RcpA active site the phosphoryl group occupies the central position and interacts with several groups as described. The sodium ion (black) is coordinated in an octahedral manner, involving two water molecules in the first coordination sphere. The side chain of S100 occupies two alternative rotameric conformations one of which (white) is averted from the phosphate group, for the price of the loss of a hydrogen bond contact (2.50 Å).

The phosphoryl moiety is involved in a set of interactions with the surrounding protein consisting of hydrogen bonds and salt bridge contacts (Fig. 4.20). These include amide groups of the peptide backbone of L71, N72 and T101, the side chain hydroxyl group of S100, T101 and the amino group of K121. Interestingly, the electron density of S100 indicates the coexistence of two alternative rotamers, one pointing towards the phosphate moiety and forming an H-bond, the other pointing away, towards the backside of the dimer interface. The putative metal binding site in the acidic surface depression was found to be occupied by a sodium ion that emulates the binding interactions, normally performed by a divalent ion like magnesium. Like in the case of  $Mg^{2+}$ -CheY [Stock, et al., 1993, Welch, et al., 1994], the first coordination sphere of this water consists of carboxylate oxygens of D18 and D70 (PHD70), the backbone carbonyl oxygen of N72, one phosphate oxygen and two additional water molecules.

#### 4.3.4.2 Apo-RcpB

The active site of RcpB represents the inactive form of the receiver. It is populated by several water molecules, one of which occupies a position, comparable to that of the metal ion in the CheY-Mg<sup>2+</sup> structure (Fig. 4.21) or the sodium ion in phospho-RcpB. This water is coordinated in a tetradedral manner by the carboxyl oxygens of D69 (2.69 Å) and D14 (2.35 Å), the backbone carbonyl oxygen of N71 (2.33 Å) and a second water (2.47 Å). Threonine 99 is rotated towards D69 and forms a hydrogen bond with the carboxyl group.

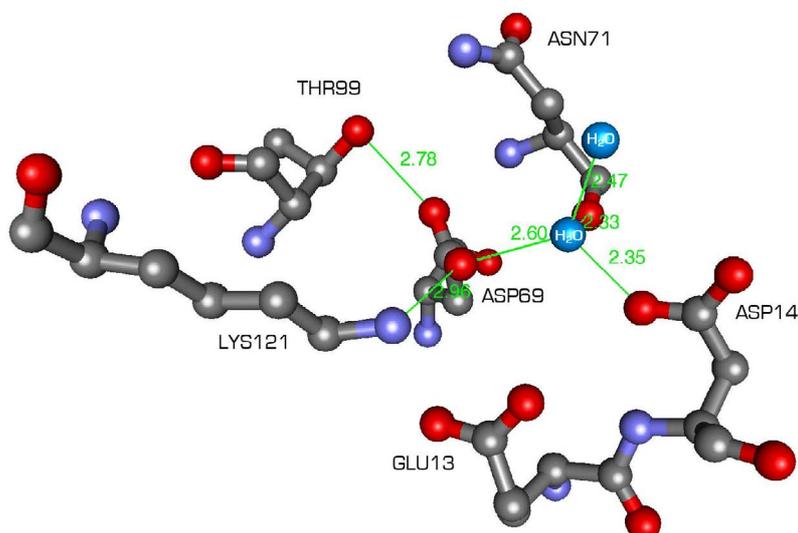


Figure 4.21. **Active site architecture in phospho-RcpA.** Representation of the conserved active site residues shown as ball and stick models. Atoms are colored according to type: C [gray], O [red], N [pale blue] and water molecules [turquoise]. In the crystal structure of RcpB the putative metal binding site was occupied by a water molecule. Its tetragonal coordination sphere is built up of the H-bonded side chain carboxyl groups of D14 and D69 and the backbone carbonyl oxygen of N71, and one additional water which tops the trigonal pyramid.

#### 4.3.5 Conformational differences between phospho-RcpA and apo-RcpB

Both proteins form the described dimer in the crystalline state and in solution. From mass spectrometric analysis it was concluded that the pure proteins, expressed in *E.*

*coli*, always were binary mixtures of both, apo- and phosphorylated form in a variable ratio. The binary samples reproducibly eluted as symmetric singular peaks from the size exclusion column with an apparent molecular mass of approx. 35 kD, which suggested a homogenous molecular size distribution.

Phospho-RcpA and apo-RcpB have a very similar active site conformation as depicted in Fig. 4.22. Except for the carboxyl group of the phosphoryl-accepting aspartate, which is rotated by  $104.4^\circ$  around the  $C_\beta$ - $C_\gamma$ -bond in phospho-RcpA, the other residues (RcpA: E17, D18, S100, K122 and RcpB: E13, D14, T99, K121) show the same rotameric orientation. The rotation of the carboxylic group of the aspartate side chain due to phosphorylation, and the presence of the bulky phosphate group in RcpA result in a lateral movement of the side chains of E17, S100 and K121, the latter being hydrogen bonded to a phosphoryl oxygen atom (2.99 Å) and the side chain oxygen of E17 (2.45 Å). This in turn leads to a slight drift of the associated backbone chain segments, especially the loop connecting  $\beta_4$  and H4 (0.84 Å) and sheet  $\beta_5$  (1.1 Å). D18, due to the water-bridged hydrogen bonding to the phosphate group, also moves approx. 0.9 Å upwards in direction to the phosphate, and thus affects the position of loop  $\beta_1$ /H1 and helix H1. The backbone atoms of the monomers of RcpA and RcpB can be superimposed with a RMS deviation of 0.99 Å.

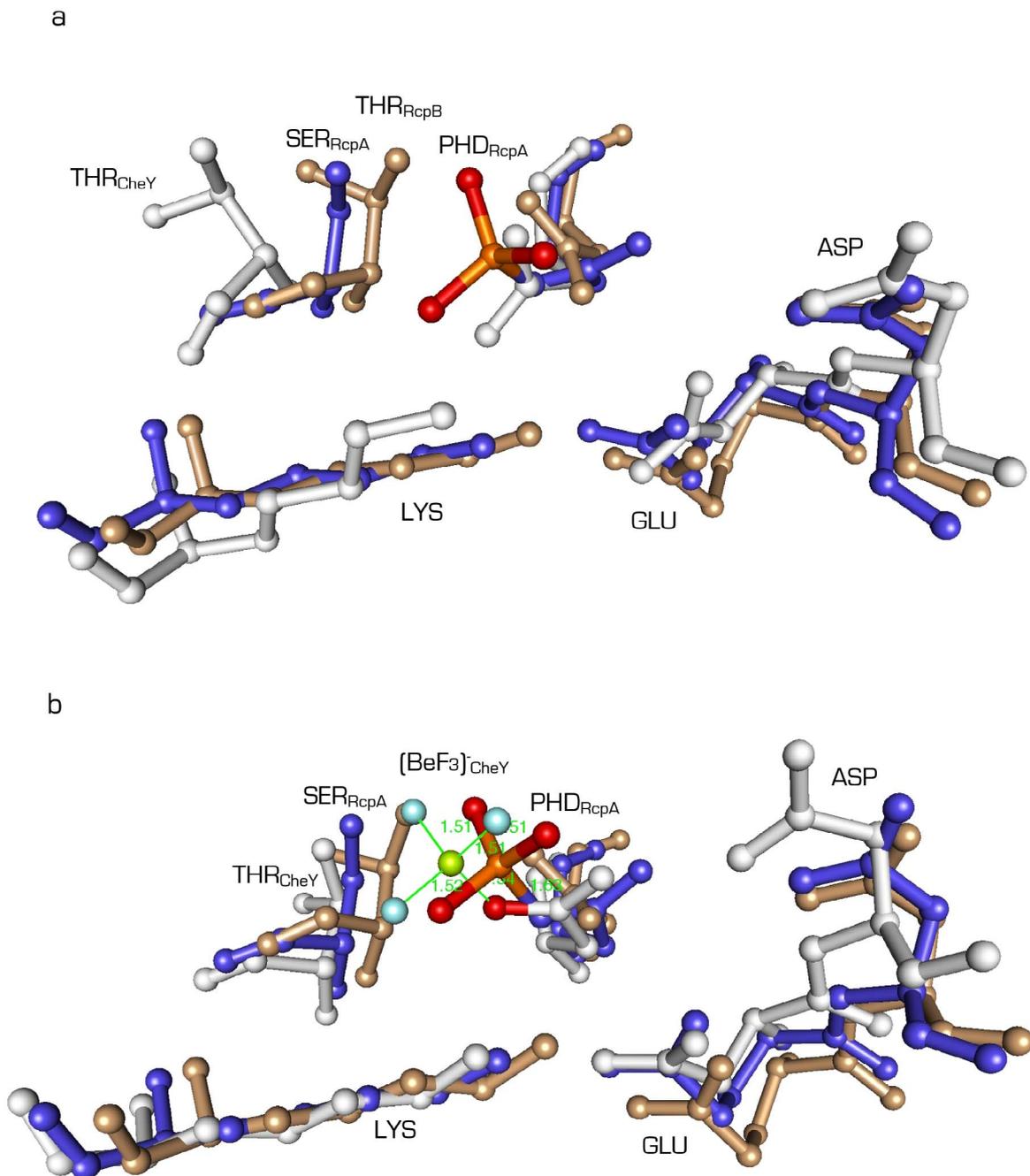


Figure 4.22. **Comparison of the active site architectures** of phospho-RcpA, RcpB and (a) CheY and (b) activated CheY with bound  $[\text{BeF}_3]$ . (a) T87 of CheY is in a remote position compared to the active state (b). In phospho-RcpA and apo-RcpB, the respective side chains (S100 and T99) occupy similar positions comparable to the active form of CheY. Interestingly, the crystallographic data show unambiguously that S100 of phospho-RcpA exists in a second alternative rotameric conformation (Fig 4.20). Whether this represents an inside glimpse into the mechanical details underlying activation is not clear. The active site of apo-RcpB resembles more the active state of CheY than the inactive state. Due to the presence of the phosphoryl group in RcpA, the side chains of K122, S100 and E17 are laterally shifted with respect to their counterparts in RcpB.

Interestingly, the side chain of S100 was found to exist in two alternative conformations in the structure of RcpB, one pointing towards the phosphate group and forming a H-bridge (2.70 Å) to one of its oxygen-atoms, the other pointing away to the H4/β5/H5 face and resides in close proximity to Y119 (distance of the hydroxyl oxygens: 3.36 Å). According to the electron density maps, both conformations have comparable occupancies although only the first is additionally stabilized by an H-bond to the phosphate-group.

#### 4.3.6 Comparison of RcpA and B with other receiver domains

The overall structures of RcpA and RcpB are similar to those of CheY homologues as anticipated from secondary structure predictions. Nevertheless, RcpA and RcpB have unique features, which discriminates them from other RRs. Primarily this comprises the intrinsic ability for dimer formation, driven by a set of highly conserved amino acid residues. Secondly, the phosphorylated form of RcpB is extraordinarily stable towards hydrolysis, which is untypical for phosphorylated aspartates in receiver modules. The structural details to the atomic level of RcpA and B enabled the identification of residues, which are involved in the formation of a specific dimeric form as described in detail in 4.3.2. BLAST searches and sequence comparison with related RRs showed generally that phytochrome associated RRs possess this motif for dimerization that is unique to a subgroup of RRs of the cheY type, namely those cotranscribed with phytochrome-like receptor histidine kinases from other bacteria.

## 4.4 The N-Terminal Receptor Domain of Plant Phytochrome A

### 4.4.1 Expression and purification of 59 and 65 kD PhyA

The heterologous expression of the N-terminal receptor domain of the Phytochrome A from *Avena sativa* was exclusively accomplished in yeast. This was done for two main reasons: first of all, yeast as a eukaryotic expression system provides the required molecular components for a correct folding and processing of a heterologous protein from a eukaryotic source organism. It is known that phytochromes expressed in *E. coli*, do not fold properly to form a functional receptor unless some certain chaperonins are co-expressed [Hill, et al., 1994]. Second of all, yeasts, constitute expression systems with certain economic vantages like easy large-scale cultivation and the ability to produce high levels of heterologous proteins.

Two fragments of the wild type phytochrome A from *A. sativa* were expressed and subjected to characterization and crystallization experiments. These fragments were chosen according to proteolytic fragments derived from trypsin digestion and comprised residues 1-595, referred to as 65 kD PhyA, and residues 66-595, referred to as 59 kD PhyA as illustrated in Fig. 4.23. In previous works, the respective coding sequences were cloned into vectors for the expression in the yeast strains *Pichia pastoris* (59 kDA PhyA) and *Hansenula polymorpha* (65 kD PhyA, Rhein Biotech, Düsseldorf) [Gärtner, et al., 1996, Mozley, et al., 1997]. The gene products comprised C-terminal polyhistidine tags for facile affinity purification.

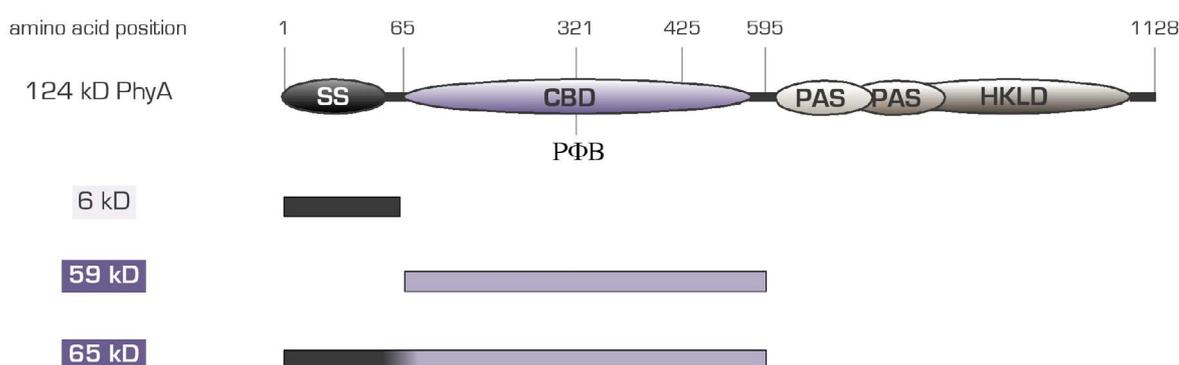


Figure 4.23. **Functional entities in *A. sativa* PhyA.** Shown is a schematic representation of the organization of main functional domains in oat PhyA. The N-terminal half of the protein contains a serine-rich stretch [SS] and the sensor domain [CBD] that binds the phytochromobilin chromophore [PΦB]. The C-terminal half harbors two Per-Arnt-Sim domains [PAS] and a histidine kinase-like domain [HKLD]. Two constructs, 59 and 65 kD PhyA, are indicated [blue]. They represent fully photoconvertible receptor domains and were selected as targets for physicochemical characterization and crystallization.

#### 4.4.1.1 Expression, *in vitro* reconstitution and purification of 59 and 65 kD PhyA

For the benefit of high cell mass production and high levels of recombinant protein, the transgene yeast strains hosting the genes of interest were cultivated by fermentation in 5 liter bioreactors as described in detail under 3.4.5. From a standard fermentation procedure, a total of approx. 500 to 1000 g wet cells were obtained. The cells were usually divided into aliquots of 50 g, flash frozen in liquid nitrogen and stored until further proceeding. For the preparation of the recombinant proteins, the cells were thawed on ice and disrupted as explained in 3.3.8. The reconstitution of the apo proteins with the desired chromophore (in general: PCB) was either done in the clarified cell lysate or after purification, as explained in 3.3.10. The purification of both fragments, 59 and 65 kD PhyA was attained *via* a two step protocol, that involved a preliminary separation from cellular proteins by metal affinity chromatography followed by a “polishing” step on a size exclusion column. The process of purification is illustrated in figure 4.24 for the 59 kD and in figure 4.25 for the 65 kD PhyA receptor domain.

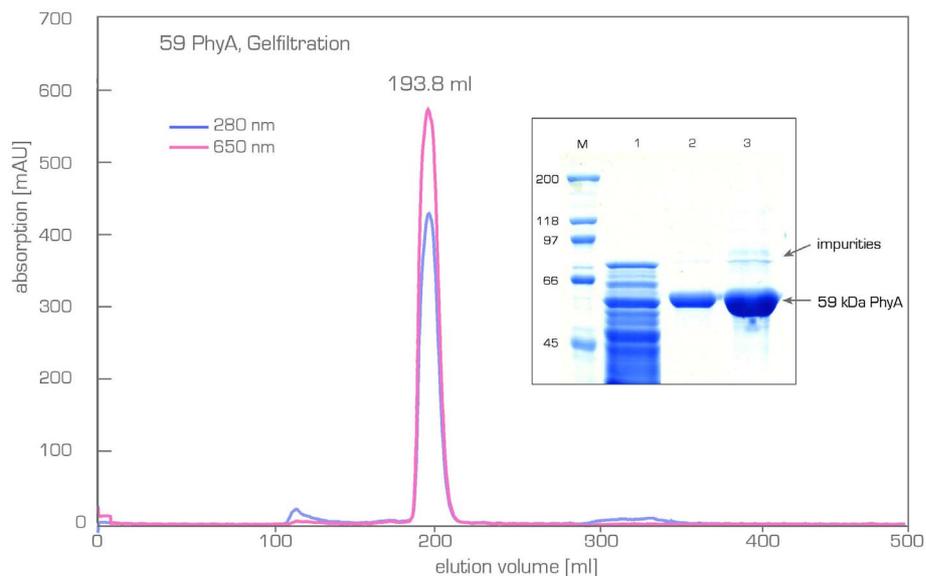


Figure 4.24. **Purification of 59 kD PhyA.** A typical size exclusion chromatogram (HiLoad 26/60 Superdex 75 prep grade column) showing the elution profile of the N-terminal receptor domain after affinity purification; flow rate 0.53 ml min<sup>-1</sup>, elution buffer G (150 mM NaCl, 5 % glycerol, 50 mM Tris-HCl pH 8.0). According to the column calibration (see 3.3.9), the elution volume of the main peak at 193.8 ml corresponds to an apparent molecular weight of approx. 82 kD.. The inset shows the coomassie stained SDS-PAGE analysis of pooled fractions in lane 2 as 20 fold and 3 as five fold dilutions. Lane 1 represents the cell extract prior to purification and lane M contains marker proteins. The denatured protein runs at the expected height. Protein samples were mixed with an equal volume of sample buffer (3.3.2), heated to 80°C for 5 min and applied to the gel. The gel was run at constant power (180 V) until the lower dye band of the sample buffer reached the lower end of the gel.

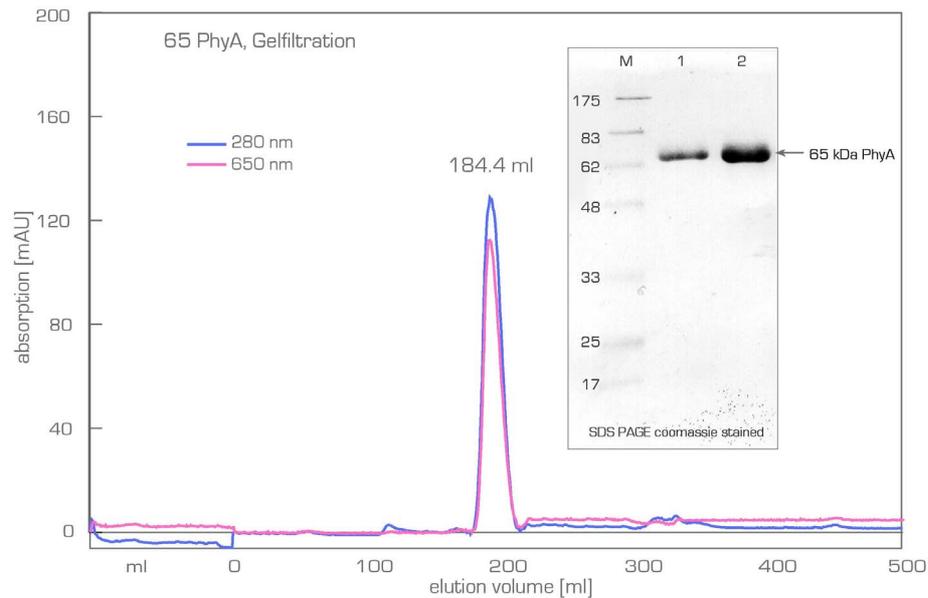


Figure 4.25. **Purification of 65 kD PhyA.** The size exclusion chromatogram (HiLoad 26/60 Superdex 75 prep grade column) shows the elution profile of the 65 kD N-terminal receptor domain after affinity purification; flow rate 0.53 ml min<sup>-1</sup>, elution buffer G [150 mM NaCl, 5 % glycerol, 50 mM Tris HCl pH 8.0]. The elution volume of the main peak at 184.4 ml corresponds to an apparent molecular weight of approx. 110 kD. The inset shows the coomassie stained SDS-PAGE analysis of the pooled fractions in lane 1 and the eluted protein from affinity chromatography in lane 2. Lane M contains the molecular weight marker. The denatured protein runs slightly above the 62 kD marker band. Protein samples were mixed with an equal volume of sample buffer [3.3.2], heated to 80°C for 5 min and applied to the gel. The gel was run at constant power (180 V) until the lower dye band of the sample buffer reached the lower end of the gel.

As the receptor domains undergo conformational changes upon absorption of light, the protein samples were illuminated with far-red light ( $\lambda > 715$  nm) to shift the photo-equilibrium between the  $P_r$  and  $P_{fr}$  states to the  $P_r$  form prior to gel filtration. The chromatography was run under dark light conditions to prevent photochemical conversion or depletion of the light sensitive chromophore.

The apparent molecular weight of the purified proteins was not consistent with the expected theoretical values for the monomeric proteins (see Table 4.5) but might represent dynamic (59 kD PhyA) or static (65 kD PhyA) dimerization equilibria. The shape and symmetry of the elution peaks suggested a high homogeneity throughout the samples.

	monomer, theor. MW [kD]	apparent MW [kD]
59 kD PhyA	60 765	82 000
65 kD PhyA	67 380	110 000

Table 4.5

#### 4.4.1.2 UV/VIS Spectroscopic characterization

The purified receptor domains were subjected to UV/VIS spectroscopical analysis in order to survey their absorption profile and photoreversibility of the photochromous proteins. Usually, spectra of the two photoreversible states,  $P_r$  and  $P_{fr}$  of all preparations were recorded and used for the calculation of the  $P_r$ - $P_{fr}$  difference spectrum. Additionally, information on the purity and concentration of the samples could be deduced from the spectra. Figure 4.26.a and b depicts typical UV/VIS spectra from highly purified receptor domains. As a measure of purity the specific absorption ration (SAR) between the  $P_r$  absorption band ( $\lambda_{max} = 654$ ) and the absorption at  $\lambda = 280$  nm is used. The SAR values for protein samples with the highest purity obtainable through the applied methods were generally in the range of 1.4 to 1.6. According to the SDS PAGE analysis, the samples were at least 95 % pure.

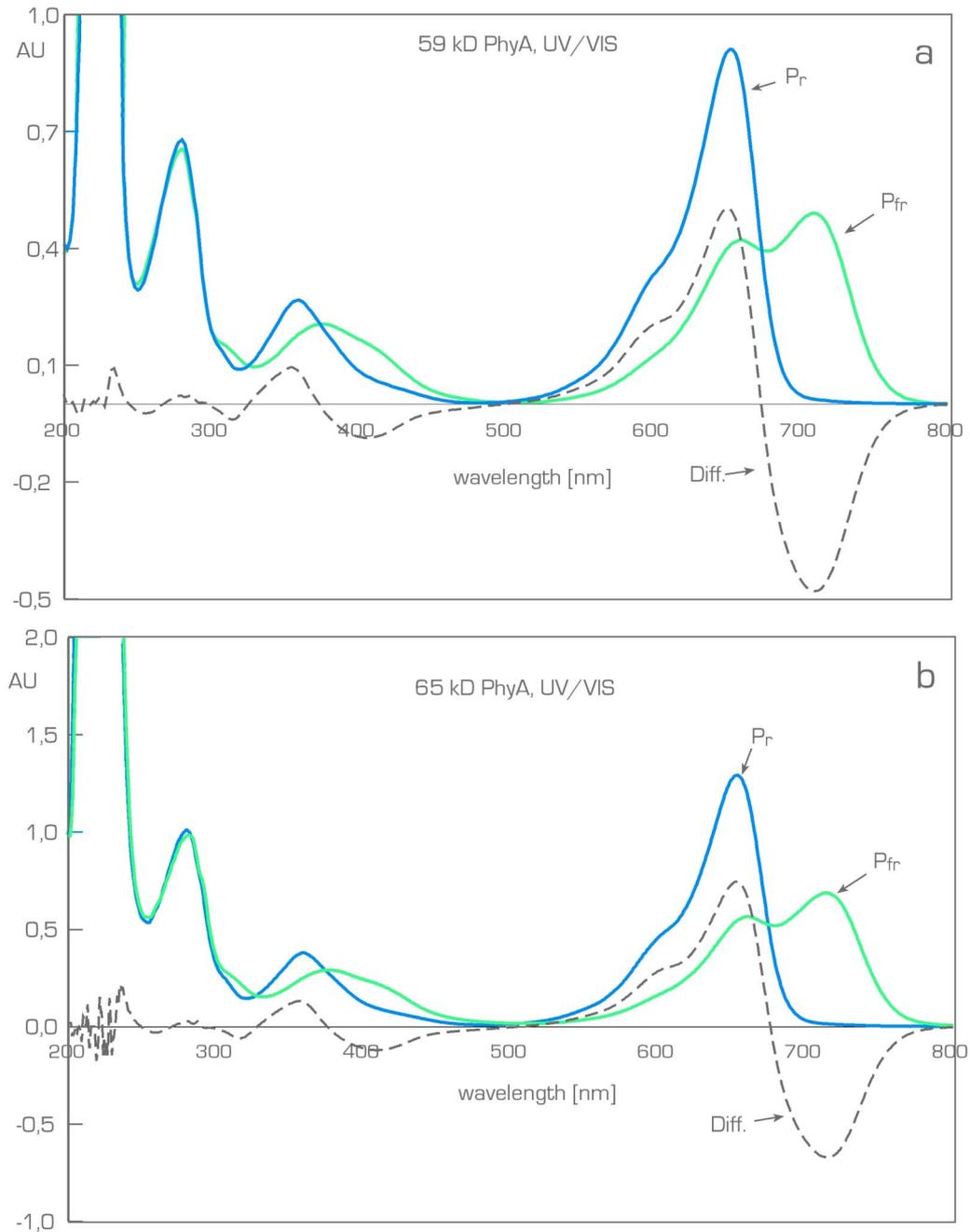


Figure 4.26. **UV/VIS Absorption spectra of 59 and 65 kD PhyA.** Depicted are typical UV/VIS spectra of (a) 59 and (b) 65 kD PhyA after the standard isolation/purification process. Both proteins exhibit the characteristic absorption band at 654 nm ( $P_r$ -form, marked blue) and the photoinducible bathochromic shift ( $\lambda_{\max} = 712$  nm) after illumination with red light ( $635 \text{ nm} \pm 8$ ), indicating the formation of the  $P_{fr}$  form (marked green). The difference spectra (dashed lines) were used to determine the protein concentration by calculating  $\Delta(\Delta A)$  between the  $P_r$  and  $P_{fr}$  absorption maxima. The spectra were recorded in 10 mm quartz cuvettes after irradiation for 60 sec with the respective light quality (red for  $P_{fr}$ , far-red for  $P_r$ ).

The concentrations of photoactive PhyA samples were calculated from the maximal deflections in the  $P_r / P_{fr}$  difference spectra at approx. 654 nm and 712 nm  $\Delta(\Delta A)$  and

employing the Lambert-Beer correlation between absorption and concentration of a given solution:

$$A = \varepsilon cd; \quad A = \log_{10}(I_0 / I)$$

The molar extinction coefficient  $\varepsilon$  for the main absorption band at  $\lambda_{\max} = 654$  of the P<sub>r</sub>-form of PhyA is 132000 [l mol<sup>-1</sup> cm<sup>-1</sup>].  $A$  is the absorbance of the sample in a quartz cuvette of path length  $d$ . From a standard purification, an average of 1.5 - 2 mg of 59 or 65 kD PhyA per gram wet cells were obtained. A summary of the determined parameters from standard sample preparation is given in table 4.6.

Sample	weight of wet cells	concentration in crude lysate	expression level	SAR after two step purification
59 kD PhyA	25	5 mg ml <sup>-1</sup>	2 mg per g cells	1.60
65 kD PhyA	25	3 mg ml <sup>-1</sup>	1.5 mg per g cells	1.50

**Table 4.6**

#### 4.4.1.3 Mass spectrometric analysis

For the detection of micro-inhomogeneities, the purified 59 and 65 kD PhyA samples were analyzed by ESI and MALTDI-TOF mass spectrometry. According to the resulting mass spectra, all samples possessed an inherent mass inhomogeneity and lacked a specific prominent signal that was in accordance with the theoretically expected masses (60765 kD for 59 kD PhyA, 67380 kD for 65 PhyA). In the range around 60 and 65 kD, respectively, the signals were broadly distributed, indicating either unspecific proteolytic activity during expression or purification, incomplete ribosomal synthesis and/or unspecific chemical modification of the protein molecules in the cell's cytosol such as phosphorylation, deamidation or oxidation.

#### 4.4.1.4 Isoelectric focussing

The results obtained from the ESI-MS analysis were further supported by the results obtained from isoelectric focussing experiments performed with both, 59 and 65 kD PhyA [outlined in 3.3.2.1]. As shown in figure 4.27, the lanes for both, 59 and 65 kD PhyA, unveil multiple bands for samples that were believed to be homogeneous. Besides that, minor differences in the distribution of the bands were observable whth in the P<sub>r</sub>- and the P<sub>fr</sub>-form of the respective fragment. As outlined in the introductory section, the receptor domains undergo conformational changes in the protein backbone, triggered

by the photoisomerization of the chromophore scaffold upon light absorption. The presence of multiple bands, differing in their isoelectric point (pI) has to be ascribed to severe inhomogeneities of the protein samples. The nature of these inhomogeneities can be of various types as discussed in section 5.2.1.

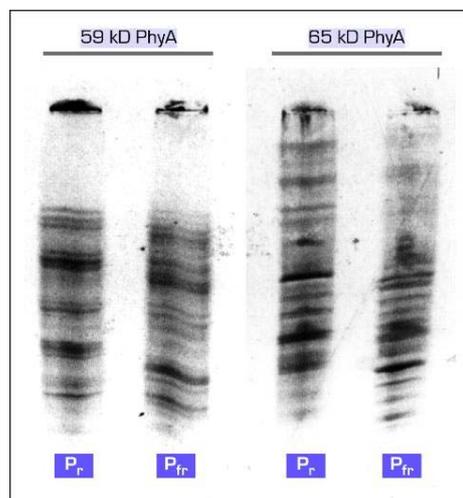


Figure 4.27. **IEF gel electrophoresis of 59 and 65 kD PhyA.** The image shows the isoelectric focussing of 59 and 65 kD PhyA in a gradient from pH 7 (top of gel) to pH 3 (bottom). The samples were shifted to the P<sub>r</sub> and P<sub>fr</sub> form and applied to the gel and separated under dark light conditions. [see 3.3.2.1].

#### 4.4.2 Circular dichroism

The purified 59 and 65 kD receptor domains of PhyA were subjected to circular dichroic (CD) spectroscopy in order to analyze the presence, quality and dynamics of secondary structural folding. As outlined in section 2.2, the two receptor fragments investigated differ by an N-terminal peptide that consists of 65 amino acids. Secondary structure prediction tools [e.g. PSIPRED, <http://bioinf.cs.ucl.ac.uk/psipred>, [Jones, 1999]] foretell that this peptide contains mostly helical structure. Published results indicate that this very N-terminus is a key player in the signal transfer event to downstream components [Fankhauser, 2002, Jordan, et al., 1996, Sweere, et al., 2001]. According to Chen and colleagues [Chen, et al., 1997], the conformational changes accompanying the  $P_r$  to  $P_{fr}$  transition involve refolding events in the N-terminus. The CD spectra were recorded as described in section 3.5.2. In brief, for the UV and near UV, the CD signals were accumulated over ten scans and averaged. Between successive scans, the respective photoisomeric form ( $P_r$  or  $P_{fr}$ ) was irradiated with the cognate light quality in order to regenerate the equilibrium state. The overlaid  $P_r/P_{fr}$  UV CD spectra of 65 and 59 kD PhyA are depicted in figure 4.28.

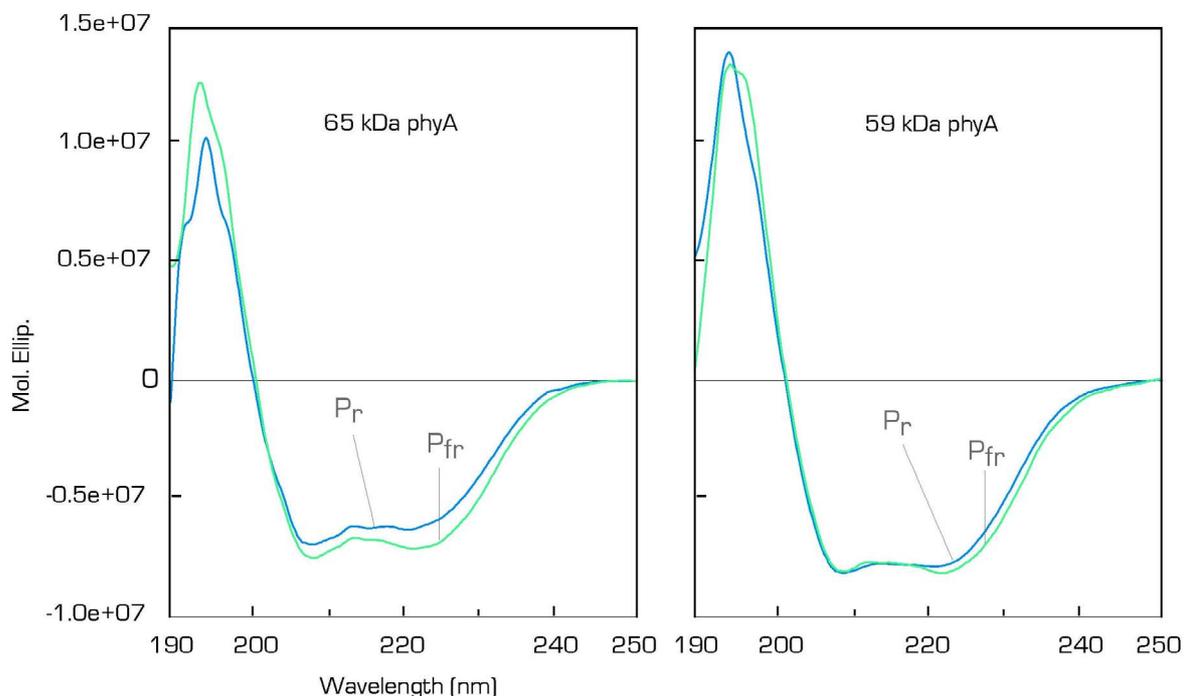


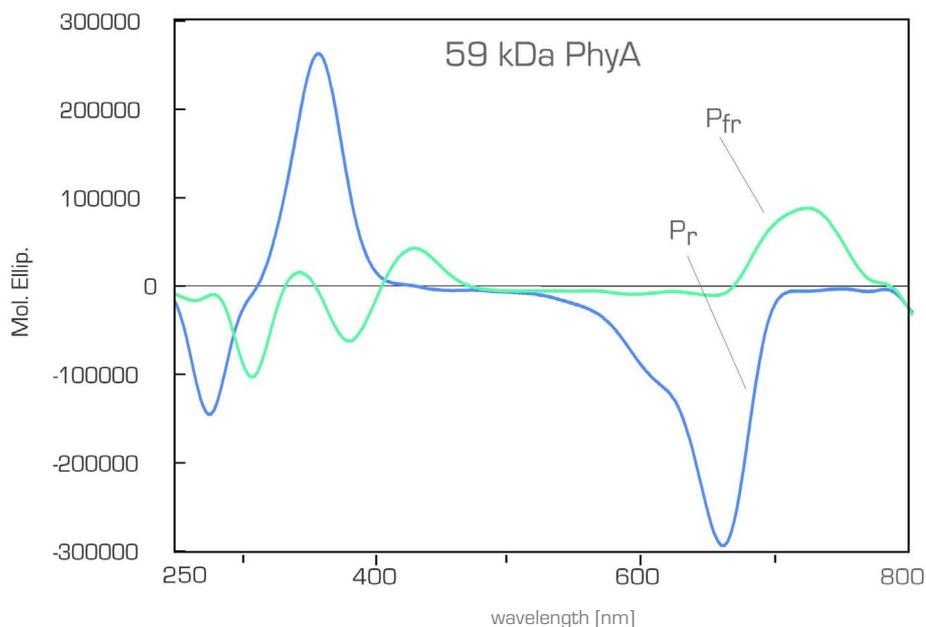
Figure 4.28. **UV-CD Spectra of 59- and 65 kD PhyA.** For both fragments, spectra of the  $P_r$  (marked blue) and  $P_{fr}$  form (marked green) were recorded and overlaid to illustrate the difference between the two photoisomeric states. The 65 kD fragment exhibits a change in secondary structure upon illumination with red light. The 59 kD fragment in turn, which lacks the 65 N-terminal amino acids, shows comparably small differences between the two forms. The spectra were recorded in 2 mm quartz cuvettes as described in 3.5.2.

Construct	$\alpha$ -helical	$\beta$ -strand	turn	random
65 kD PhyA Pr	22.2	42.3	11.1	24.3
65 kD PhyA Pfr	30.6	34.6	11.2	23.7
59 kD PhyA Pr	28.3	36.8	12.2	22.6
59 kD PhyA Pfr	29.2	31.2	18.2	21.4

Table 4.7. **Secondary structure elements** in 65- and 59 kD PhyA in the P<sub>r</sub> and P<sub>fr</sub> form listed as percent [%]  $\alpha$ -helix,  $\beta$ -strand, turn and random coil., calculated with the program DICRPROT [Deleage and Geourjon, 1993], using the least-square-fit routine.

The changes in the CD spectra upon conversion of the red to the far-red absorbing form are fully reversible as could be demonstrated for both fragments (data not shown). Also the recycled P<sub>r</sub> CD signal almost completely reverts to the initial P<sub>r</sub> signal, which implies the reversibility of the unfolding event.

CD spectra recorded in the visible region exhibit the CD signal resulting from the small molecule chromophore PCB incorporated in the protein scaffold. The opposite rotation of the polarized light after passing through the P<sub>r</sub> and P<sub>fr</sub> samples, respectively, reflected the photo-induced rearrangement of the tetrapyrrole upon light absorption. The P<sub>r</sub> CD spectrum is very similar to the characteristic P<sub>r</sub> UV/VIS spectrum.



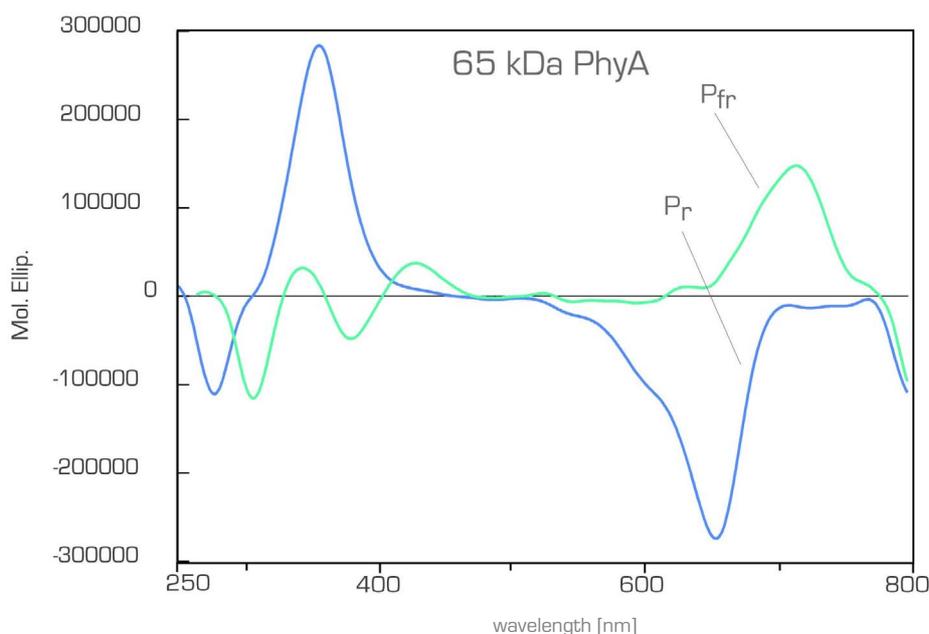


Figure 4.29. **VIS-CD Spectra of 59 and 65 kD PhyA.** CD effect of the protein-incorporated phycocyanobilin chromophore, measured between 250 and 800 nm for  $P_r$  (marked blue) and  $P_{fr}$  (marked green). The spectra were recorded in 10 mm quartz cuvettes and represent accumulations of ten successive scans. After each scan, the samples were irradiated for 30 sec with light of the initial wavelength range ( $\lambda = 635 \pm 8$  nm for  $P_{fr}$ ,  $\lambda > 710$  nm for  $P_r$ ) to maintain the equilibrium state of the respective state ( $P_r$  and  $P_{fr}$ ).

#### 4.4.3 NMR spectroscopic characterization of recombinant phytochromes

NMR offers a sophisticated technique for monitoring the chemical surrounding of a given nucleus. For example, classic  $^1\text{H}$ -NMR supplies information on the chemical context of (non-acidic) protons in a molecule by giving specific nuclear magnetic resonance signals. A different, two-dimensional NMR technique,  $^1\text{H}$ - $^{15}\text{N}$ -HSQC-NMR is of special value for the characterization of biomolecular probes such as proteins. The principles of this technique are outlined in section 3.5.3. HSQC spectra show resonance signals for each amide group in a given protein. The chemical shift depends upon the local environment of the amide group, i.e. the type of secondary structure. Due to the low natural abundance of the NMR active nitrogen isotope  $^{15}\text{N}$ , the complete labeling of the macromolecular probe with  $^{15}\text{N}$  is a necessity. In the case of PhyA, this was accomplished by expressing the protein in yeast, growing on a minimal medium with an inorganic compound, such as  $^{15}\text{N}$ -ammonium chloride as the sole source of nitrogen. As NMR usually requires high amounts of proteins (concentrations of at least 0.1 mM), the yeast strain was grown by fermentation according to the procedure described in 3.4.5. A technical problem that was experienced during fermentation was the growth limiting effect caused by a shortcoming of the  $^{15}\text{N}$ -ammonium salt in the medium. The cell growth ceased (monitored by the  $\text{OD}_{600}$  value of the culture and uptake of base for

maintaining the initial pH in the medium] as soon as the nitrogen source was consumed and restarted when it was replenished. In order to optimize the conditions for the fermentation, test runs were performed with regular ammonium chloride instead of the  $^{15}\text{N}$ -labeled homologue. The expression of both domains, 59 and 65 kDa PhyA, was monitored by Western blot analysis, during 24 and 29 hrs after induction. The recombinant proteins were detected with an antibody specific for the His<sub>6</sub> peptide. Figure 4.30 shows the chemoluminographs of the optimized test expressions. The expression levels reached a maximum approx. 20 hrs after methanol induction. Also detectable in the blots were His<sub>6</sub> tagged protein fragments with a higher electrophoretic mobility, that represent accumulated proteolytic fragments.

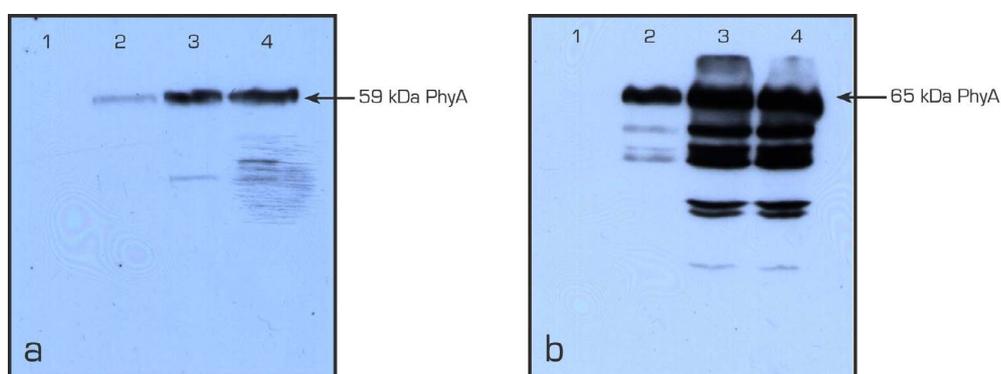


Figure 4.30. **Western Blot analysis of expression rates in  $^{15}\text{N}$ -labeling experiments.** Illustrated are the relative expression rates of (a) 59 and (b) 65 kD PhyA in *P. pastoris* cultures grown by fermentation on minimal medium containing  $^{15}\text{N}$ -ammonium chloride as described in 3.5.4. The chemoluminographs show the immuno-detected, His<sub>6</sub> tag bearing proteins. In the case of 59 kD PhyA, samples were taken prior to methanol induction (lane 1) and after 4 (lane 2), 10 (lane 3) and 24 hrs (lane 4) of methanol feeding. For the 65 kD fragment, samples were taken after 5 (lane 2), 21.5 (lane 3) and 29 hrs (lane 4). The protein extracts were prepared as described [see paragraph 3.4.4] and separated by standard SDS PAGE.

Fermentation in the presence of  $^{15}\text{N}$  ammonium chloride was performed in 2 liters initial medium and yielded approx. 200 g wet cells per liter. In the case of 59 kD PhyA, one gram of wet cells contained about 0.4 mg of the expressed receptor domain as determined from reconstitution experiments. The  $^{15}\text{N}$ -labeled 59 kD recombinant protein was purified and reconstituted with PCB as described, and characterized by UV/VIS spectroscopy. The final preparation exhibited the phytochrome characteristic spectrum and possessed a high purity that was similar to that of unlabeled preparations. The concentration of the sample was elevated to 0.1 mM and the probe subjected to NMR spectroscopy, using a Bruker Avance DMX 750 MHz NMR spectrometer in Garching (TU Munich) for the  $^1\text{H}$ - $^{15}\text{N}$  HSQC experiments and a Bruker 600 MHz NMR for standard

$^1\text{H}$  spectra. The latter were recorded in order to check whether the sample concentrations and purity were suitable for the HSQC experiment. Figure 4.31 shows a typical  $^1\text{H}$ -NMR spectrum from 59 kD PhyA.

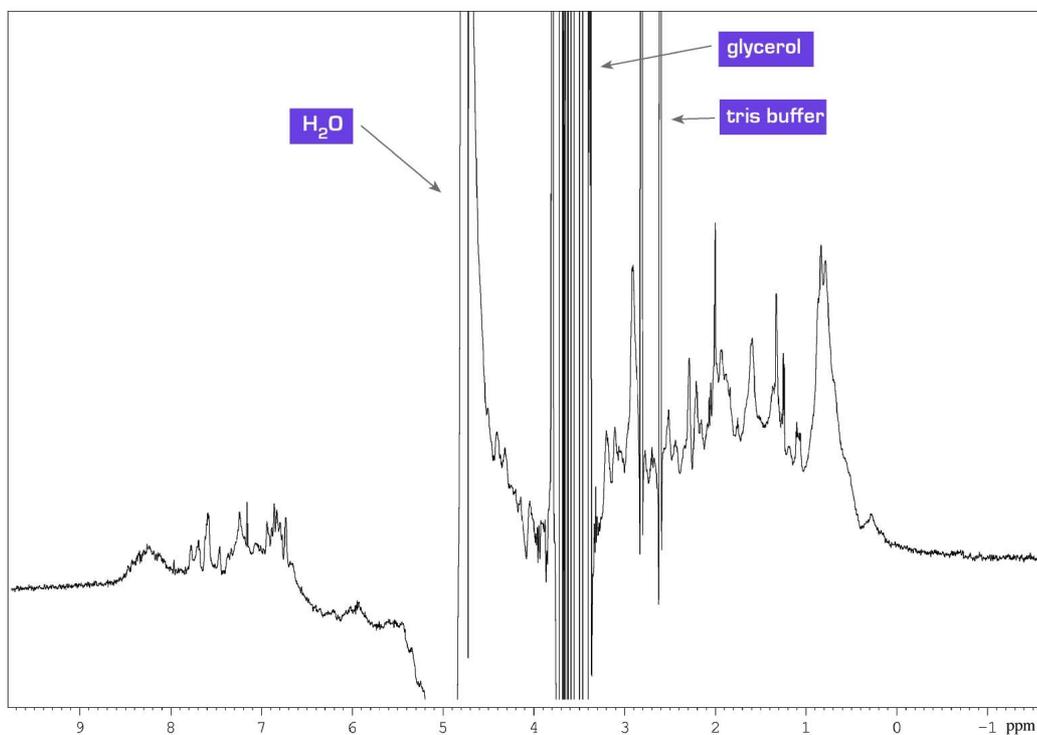


Figure 4.31.  $^1\text{H}$  NMR of 59 kD PhyA. Shown is a typical one-dimensional proton NMR from the  $P_r$  form of a 59 kD PhyA sample. The concentration of the protein was 0.1 mM in a glycerol containing tris buffer. The spectrum was recorded at 600 MHz. The characteristic resonances of amide protons (8.0 - 8.5 ppm) show no significant dispersion.

For the two-dimensional heteronuclear experiment, the resonance signals were measured by accumulating multiple scans over-night at room temperature. Figure 4.32 depicts the  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra of the  $P_r$  and  $P_{fr}$  state of the 59 kD fragment.

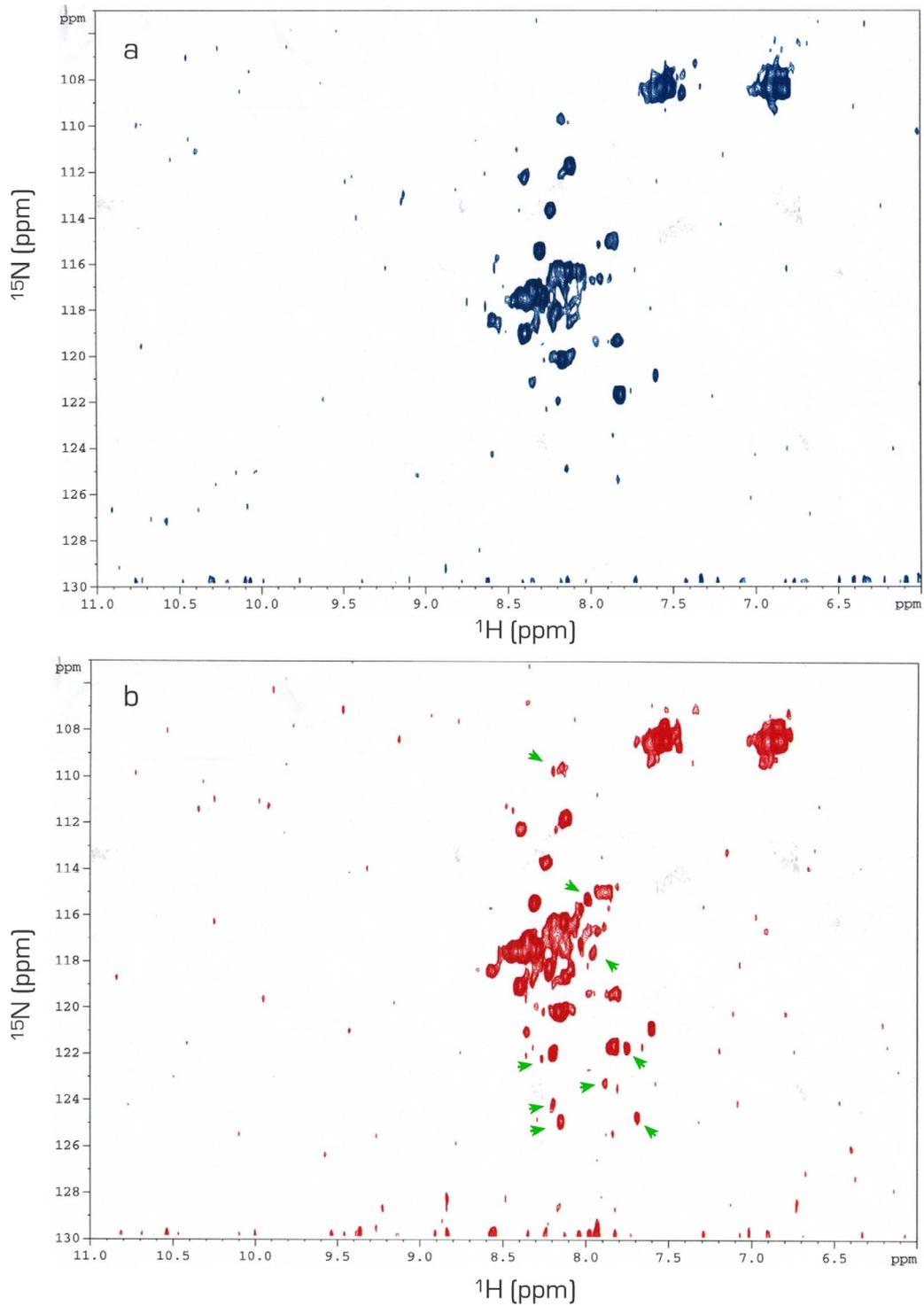


Figure 4.32.  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra of uniformly labeled 59 kD PhyA. The protein concentration was approx. 0.1 mM and spectra were obtained at 750 MHz after shifting the  $P_r/P_{fr}$  equilibrium to the (a)  $P_r$  form (irradiation with far red light) and (b)  $P_{fr}$  form (irradiation with red light). In both spectra, the amide resonances show no significant shift dispersion beyond that of a random coil protein (8.0-8.5 ppm). Differences between the two photoisomers are of minor nature and indicated with arrows. Most of the theoretically expected amide resonances are absent or significantly broadened.

#### 4.4.4 Crystallization screening

The aim for exploring a protein's structure to the atomic level by X-ray crystallography requires the well ordered, three dimensional, crystalline arrangement of the molecule in question. In some cases, this proves a tedious task to perform. In the case of PhyA, both fragments were extensively tested with regards to their suitability for single crystal formation. For this to be accomplished, the purified and spectroscopically characterized proteins were incubated with standard crystallization conditions as generally described in section 3.6.1. These conditions represent a sparse matrix approach (Jancarik and Kim, 1991), covering a large space of well distributed conditions and generally supply a useful starting point for the development and improvement of optimized conditions for crystal growth. The 59 and 65 kD fragments were tested with Crystal Screen I, II and Cryo Screen (Hampton Research), altogether comprising 150 conditions. In extended screenings, combinations of these initial conditions with a diversity of additives listed below were tested. The crystallization experiments were set up under inert green light, at room temperature in a dark room. Prior to pipetting and incubation in a thermostatic cabinet at 18°C, the probes were illuminated with either red (635 nm ±8) or far-red light (>715 nm). The crystallization droplets contained PhyA protein solution with variable concentrations and were joined with precipitating conditions in different ratios, covering the range between 3:1 (three µl protein solution combined with one µl precipitant solution) and 1:2. Because of the high tendency of both fragments to precipitate, the crystallization conditions were additionally diluted in a one to one ratio with water. The droplets were inspected under a microscope, equipped with a green-light filter for inert illumination. This was done immediately after the crystallization was completed, and subsequently each day for one week. Other crystallization setups were inspected only after two or three weeks after pipetting, in order to keep the droplets unperturbed, for as long as possible. In some cases, the droplets were observed for as long as six months.

Additives tested in the 150 basic conditions:

- glycerol (up to 20 %)
- 2-mercaptoethanol (1 mM)
- Triton X-100 (up to 0.2 mM); CMC = 0.20 mM
- n-octyl glucoside (up to 14.5 mM); CMC = 20-25 mM
- C8E4 (tetraoxyethylene n-octyl ether); CMC = 7.0 mM
- Detergent Screen, Hampton Research

None of the listed additives meant an improvement in terms of crystallogenesis. Glycerol, which generally has a beneficial effect on proteins in solution (Sousa, 1995),

could slightly retard the precipitation. The solubility of the phytochrome fragments was noticeably increased in the presence of glycerol. The addition of detergents was fatal for the proteins, in that they caused immediate precipitation of denatured protein even at concentrations below the CMC (critical micellar concentration). Detergents provided with the Detergent Screen (Hampton Research) were initially tested for their effect by simply adding them at final concentrations of one tenth of the CMC concentration to protein solutions used in crystallization.

If the phytochrome concentration was at elevated levels ( $> 4 \text{ mg ml}^{-1}$ ), the protein precipitated in every (!) condition tested. Droplets, where the protein concentration was below that, remained clear in some cases even after several months, indicating that supersaturation was not achieved in the respective condition. A total of more than 5000 different conditions were tested with both phytochrome fragments, but crystal growth was not attained.

#### 4.5 Cloning and expression of phytochrome interacting factor 3

Phytochrome interacting factor 3 (PIF3) from *Arabidopsis thaliana* is a 57 kD soluble protein built up of 524 amino acids that interacts with phytochrome A and B as referenced in the introductory chapter (2.1.4).

The gene encoding for *pif3* was amplified by RT-PCR from total mRNA preparations and cloned into the vector pET-28a(+) between the NdeI and XhoI restrictions sites as described in 3.2.19. For heterologous expression in *E. coli*, the recombinant plasmid was transformed into the strain BL21(DE) Gold and positive transformants were selected on kanamycin. A preliminary test expression according to 3.3.1 at 30°C readily yielded a protein of the expected size, but exclusively in inclusion bodies. Modified test expressions were performed at lower temperatures (15°C for 24 hrs or at 24°C for 3 hrs) and induction with 100 and 400  $\mu\text{M}$  IPTG. Even under these moderate conditions, PIF3 was only produced in insoluble form as inclusion bodies. Figure 4.33 shows the SDS PAGE analysis of a test expression at 15°C.

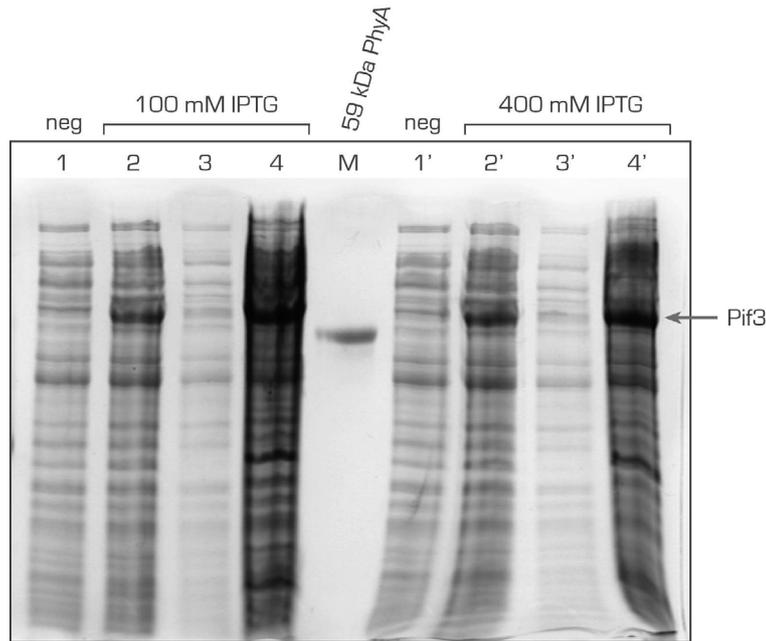


Figure 4.33. **Test expression of PIF3 in *E. coli*.** Shown is the coomassie stained SDS PAGE analysis of *E. coli* cell extracts from a culture expressing the *PIF3* gene. The expression was induced with either 100 or 400  $\mu$ M IPTG and continued for 24 hrs at 15°C. The lanes contain samples of the denatured whole-cell extract taken immediately prior to induction (1 and 1') and after 24 hrs (2 and 2'). Lines 3 and 3' contain the clarified soluble (3 and 3') and 4 and 4' the insoluble fraction after cell lysis, M contains 59 kD PhyA as a marker protein. PIF3 (57 kD) was expressed in detectable amounts but the major part resided in the insoluble fraction. The soluble part contained no significant amounts of PIF3 as judged from the gel.

#### 4.6 Cloning of *A. thaliana* PHYA, PHYB and PIF3

The coding sequences of the full length receptor proteins PhyA and PhyB, and the interacting factor PIF3 from *Arabidopsis thaliana* were cloned into the vector pPICZ A for expression in yeast (*P. pastoris*). As described in 4.5, *E. coli* proved to be no suitable system for the expression of PIF3. The genes for all three proteins were amplified from a *A. thaliana* cDNA library with the respective primers as described in 3.2.19-20 (Fig. 4.34). The PCR primers were designed to introduce the respective 5' and 3' endonuclease recognition sites, a 5' yeast consensus sequence for transcription initiation and the coding sequence for a C-terminal thrombin cleavage site upstream to the C-terminal His<sub>6</sub> tag provided with the plasmid [see 3.2.19-20 and table 4.8 for details].

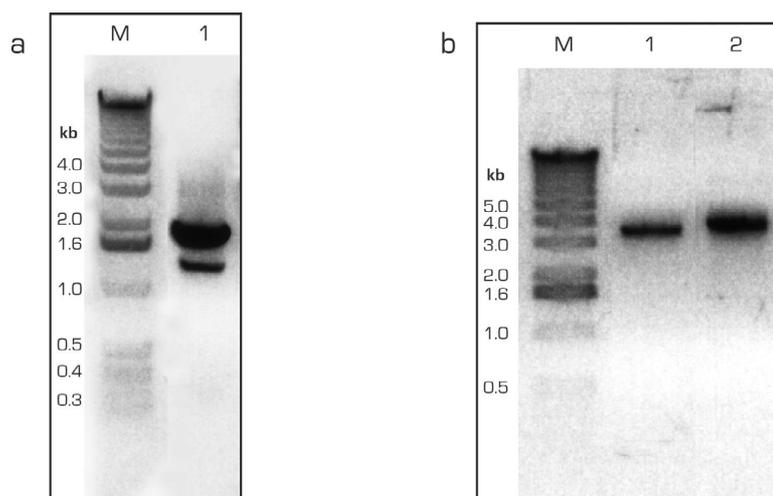


Figure 4.34. **PCR amplification of PIF3, PHYA and PHYB genes from *A. thaliana*.** The agarose gel electrophoretic analysis shows the ethidium bromide stained DNA fragments of (a) PIF3 [lane 1] and (b) PHYA [lane 1] and PHYB [lane 2], amplified with Pfu polymerase from an *A. thaliana* cDNA library. Lanes M contain a DNA molecular weight standard; numbers refer to kilo bases.

The yeast consensus sequence for initiation of transcription, 5'-AYA ATA ATG TCT-3' (Y = C/T) substitutes the initial ATG in the phytochrome genes. Restriction digestion and ligation were performed as explained in 3.2.6-9, and recombined plasmids were selected and propagated in the *E. coli* strain DH5 $\alpha$ . The correctness of the cloned genes was controlled by sequence analysis.

	5'-Restriction site	3'-Restriction site	Yeast consensus	Thrombin site	Final plasmid
PIF3	EcoRI	NotI	+	+	pPICZ Pfi3
PhyA	PmlI	NotI	+	+	pPICZ PhyA
PhyB	SfiI	NotI	+	+	pPICZ PhyB

**Table 4.8**

## 5 Discussion

### 5.1 Crystal Structures of RcpA and RcpB

Response regulators (RRs) of the bacterial two-component signaling system represent the classic paradigm of a molecular switch in prokaryotes and constitute central players in the cell's permanent responsiveness to the everchanging surrounding conditions. In the modular two-component pathways, RRs serve as the receiver (REC) for the information transmitted by a histidine kinase (HK) (Parkinson and Blair, 1993, Parkinson and Kofoid, 1992). The informational bit, that is handed over between the two-component elements, is a phosphate group, covalently bound to a highly conserved histidine (HK) or aspartate (RR) residue. Over the past several years, considerable efforts have been made to understand the underlying molecular mechanisms by which the comparably small RR molecules accomplish their different functions in the various scenarios of bacterial sensing, covering chemotaxis, phototaxis or osmosensing (for reviews see (Buckler, et al., 2000, Foussard, et al., 2001, Stock, et al., 2000, Stock and Da Re, 2000)). The structural data on several inactive and active (phosphorylated) representatives of the response regulator superfamily, that has been gathered over the last years, has provided a comprehensive, yet confusing picture of the signaling mechanism and it seems that every new piece of information that is published becomes a new piece of an unfinished puzzle. Apparently, this is also true for the crystal structures of two cyanobacterial, phytochrome associated RRs, RcpA and RcpB from the cyanobacterium *Calothrix* sp. PCC7601, which seem to constitute yet another new variation to a well known theme.

#### 5.1.1 Function and activation of the chemotaxis receiver domain CheY

The *E. coli* chemotaxis signaling system (Fig. 5.1) comprises one of the best understood cellular sensory-response networks and ideally serves as a general reference base for the comparison and description of related systems. The ability of *E. coli* and many other microorganisms to respond to the chemical composition of the surrounding environment is attributed to a set of interacting proteins that form a pathway, along which the initial signal is transferred by a specific phosphorelay (Falke, et al., 1997) as illustrated in Fig. 5.1.

At the very beginning of this pathway acts a receptor-transducer protein (MCP, methyl accepting chemotaxis protein), that specifically binds a small molecule ligand (repellent or attractant) on the outside of the cell and transduces this event through the membrane to the cytosolic side. The next component in the pathway is the cytoplasmic

protein CheA, a dimeric class II histidine kinase [Dutta, et al., 1999], which interacts with the membrane receptor and cross-autophosphorylates when the receptor binds a repellent [Grebe and Stock, 1999, Liu, et al., 1997].

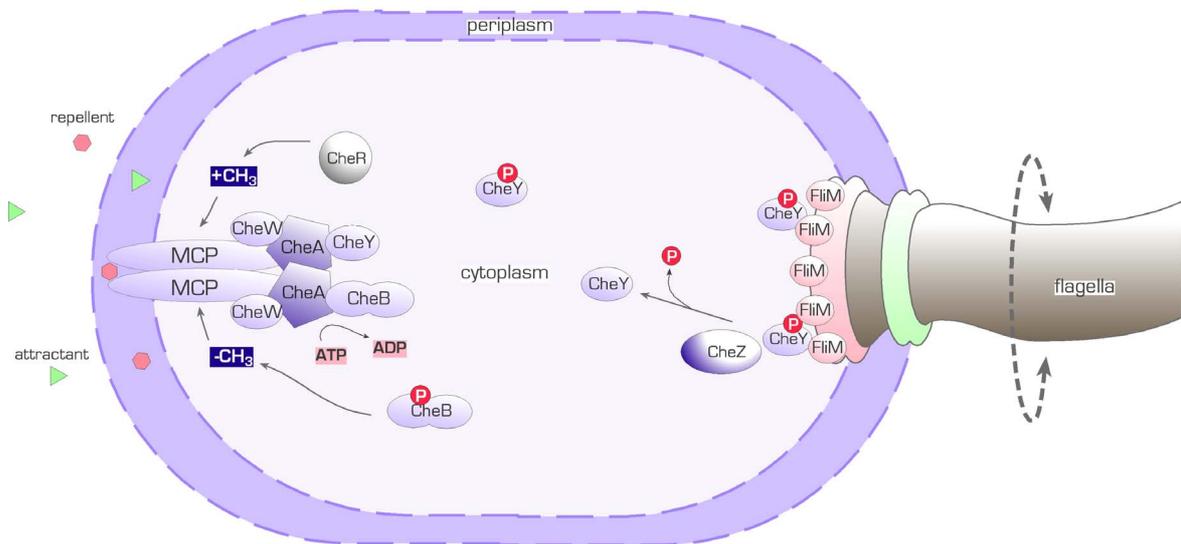


Figure 5.1. **Chemotaxis network in *E. coli***. The binding of a repellent molecule by a receptor (MCP) results in the activation of CheA kinase. CheA autophosphorylates using ATP and binds CheY to transfer a phosphate group to a conserved aspartate. Activated (phosphorylated) CheY dissociates from CheA and binds to FliM, a component of the flagellar motor. High levels of phospho-CheY increase the switching probability of the motor leading to a tumbling swimming behavior. CheZ is the CheY specific phosphatase and deactivates CheY. The methyl-erasure CheB and the methyltransferase CheR are involved in adaptation of the response network.

Autophosphorylation takes place at a highly conserved histidine residue in a structurally conserved motif called P1 domain and by using ATP as the phosphodonor. This phosphate group is then passed to an aspartic acid residue in the receiver CheY, which triggers the signal output [McEvoy, et al., 1999]. Alternatively, the phosphate group can be transferred to an aspartate in the methyl-erasure CheB for adaptation. Activated (phosphorylated) CheA interacts with CheY via an intimate interaction that is established mainly between the so called P2 domain of CheA and the H4- $\beta$ 5-H5 surface of CheY. Figures 5.2.a and 5.2.b show the crystal structure of the binary complex between the recognition domain of CheA (residues 124-257) and CheY. The buried surface covers an area of 1200  $\text{\AA}^2$  and is highly complementary [surface shape complementarity  $S_c = 0.74$ , [Lawrence and Colman, 1993]].

Transfer of the phosphate group from the P1-domain histidine in CheA to the active site aspartate (D57) in CheY is a magnesium-dependent step, like in all two-component

systems [Stock, et al., 1993], and results in a decreased affinity of CheY for the binding of CheA [Morrison and Parkinson, 1994, Stewart, 1997] and increased affinity for its target FlIM [Welch, et al., 1994], a component of the flagellar motor switch [Blair, 1995]. Figures 5.2.c and 5.2.d illustrate the molecular structure of  $[\text{BeF}_3^-]$  activated CheY in complex with an N-terminal peptide of FlIM (16N-FlIM) [Lee, et al., 2001].

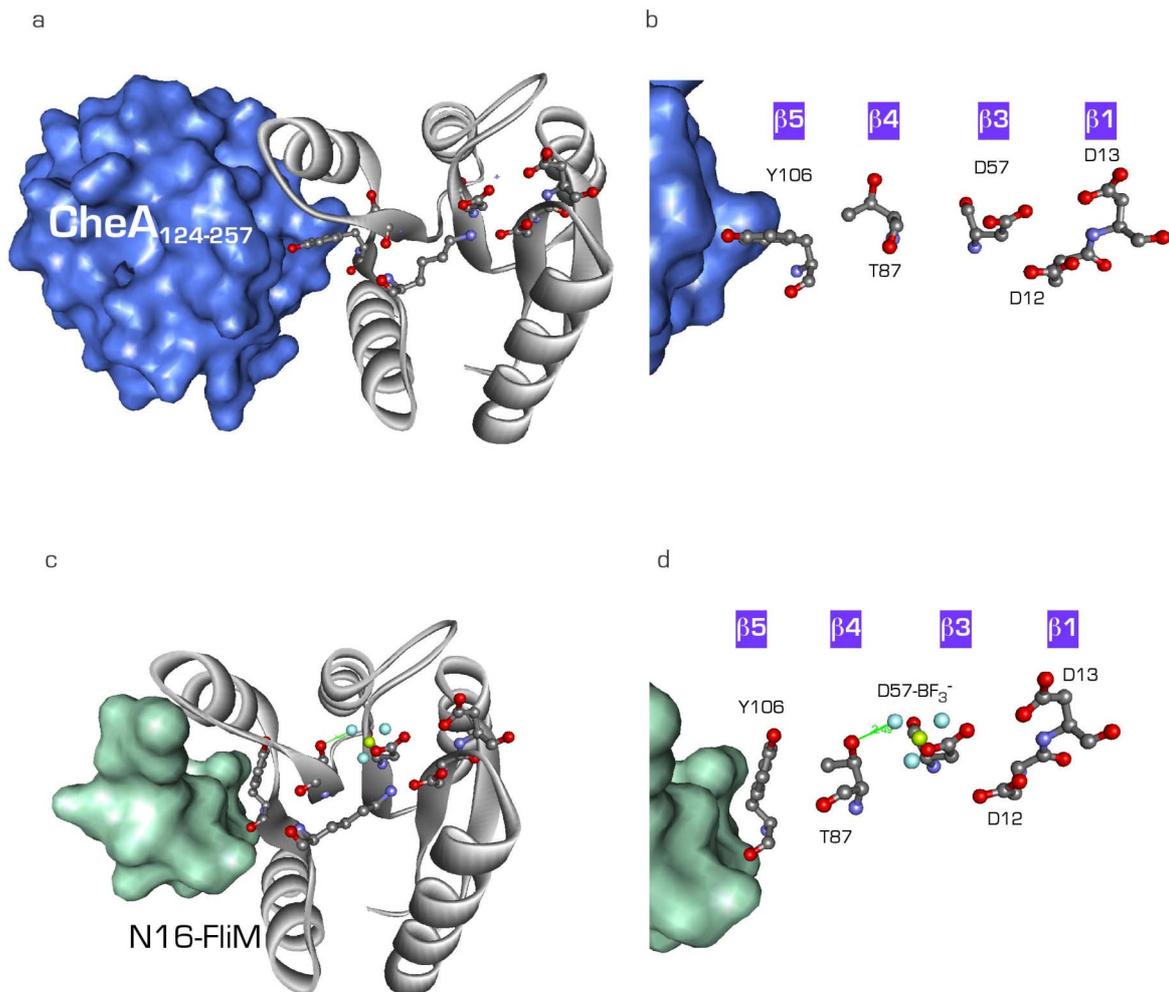


Figure 5.2. **Inactive and active CheY.** The illustration shows ribbon representations of CheY (marked silver) in (a) inactive and (b) active form with conserved active site residues as ball and stick models. The interacting partners (a) CheA<sub>124-257</sub> and (b) 16N-FlIM bound to the H4-β5-H5 surface of CheY are represented by their solvent accessible molecular surfaces. The inactive state of CheY binds activated (phosphorylated) CheA and inherits its phosphate group. The resulting activation of CheY leads to a flip of T87, rotation of Y106 and a higher affinity for binding FlIM, the event that finally causes a flagellar switch. The conformational state of side chains involved in the activation event are depicted in (c) for inactive and (d) for active CheY, together with flags, indicating their dedicated secondary structure elements [Lee, et al., 2001, Welch, et al., 1998].

This peptide contains most of the binding determinants of full length FliM (Bren and Eisenbach, 1998) and forms a helical entity that packs against the H4- $\beta$ 5-H5 face. In *E. coli*, binding of activated CheY to FliM reverses the direction of flagellar rotation from counter clockwise (smooth swimming) to clockwise (tumbling motion) (Falke, et al., 1997). The fate of activated CheY after fulfilling its signaling task, is to be dephosphorylated, autocatalytically or with assistance of the phosphatase CheZ, and to reenter the signaling pathway. On the atomic level, switching CheY to the *on* state by phosphorylation entails the rearrangement of specific residues, that are highly conserved in the response regulator superfamily. As illustrated in Fig. 5.2, the side chain of T87 of  $\beta$ 4 in active CheY (Fig. 5.2.d) is flipped towards the aspartate associated [BeF<sub>3</sub>] complex and forms a hydrogen bond with a fluoride ion. In inactive CheY (Fig. 5.2.b), the T87 side chain resides in a more remote position with the methyl group positioned in a hydrophobic pocket created by  $\beta$ 4, H4 and the intervening loop ( $\beta$ 4-H4-pocket) (Ganguli, et al., 1995, Zhu, et al., 1996, Zhu, et al., 1997). The consequence of the threonine flip is the freeing of space within the hydrophobic pocket which in turn enables the side chain of Y106 in  $\beta$ 5 to move from its solvent exposed position in the center of the H4- $\beta$ 5-H5 spanned surface to a more buried one. Thus, the conformation of Y106 is coupled *via* T87 to the phosphorylation state of active site D57 (Y-T coupling). The altered physicochemical and sterical properties of the H4- $\beta$ 5-H5 surface lead to a lowered affinity toward CheA and a higher affinity toward FliM.

The X-ray structural analysis of phospho-RcpA and apo-RcpB revealed another, slightly altered picture. Primarily, both cyanobacterial response regulators seem to be constitutively dimeric and dimerization occurs through tight interaction of the H4- $\beta$ 5-H5 surfaces of each protomer, as described in 4.3.2. A similar interaction has not yet been observed in other receiver domains and represents a new feature in the RR family. Secondly, the inactive form of the receiver RcpB exhibits structural features, that are similar to the active form of CheY, a finding, that does not correlate with the signaling model as described above and will therefore be discussed in more detail in the following.

### 5.1.2 Dimerization of RcpA and RcpB

The results described in 4.1 and 4.2 show unequivocally that the RcpA and B probes purified from *E. coli* lysates were always composed of binary mixtures of apo- and phosphorylated Rcp. The fact, that the heterologously expressed proteins were obviously phosphorylated in an unspecific way, is not as surprising as the bare fact, that it was possible to isolate the phosphorylated form in the first place. The former has been described previously (McCleary and Stock, 1994) and is probably due to small molecule phosphodonors, e. g. acetyl phosphate which is constitutively present in *E. coli* cells. The

latter, namely the extraordinary stability of the phosphorylated Rcps was unexpected. Both phospho-Rcps withstood the purification procedure and could be detected by mass spectrometric methods even days after isolation. One intrinsic feature of the aspartyl phosphate anhydride is its lability towards hydrolysis, and the lifetimes of phosphorylated CheY and other receiver domains lie in the range of seconds or minutes, even in the absence of a specific phosphatase [like CheA in *E. coli* chemotaxis], a fact that has always prevented the crystallization of CheY in its phosphorylated form. The lifetime of phosphorylated RcpA and B obviously surpasses that of CheY by far.

The most remarkable difference between RcpA and B and other structurally characterized receiver domains (figure 5.3) is their intrinsic ability for homodimer formation. Although this was indicated by results from gelfiltration chromatography experiments (4.1.1, 4.1.2 and 4.2.1), the exact nature of the interaction was only elucidated by investigating the crystal structures of RcpA and B. This capability can be ascribed to a set of highly conserved residues, distributed over the surface, that is formed by helix 4,  $\beta$ -sheet 5 and helix 5 (see results, section 4.3.2). An additional contribution arises from residues, located in the very C-terminal parts of RcpA and RcpB. The extended (as compared to CheY) C-terminus is a conserved feature in a subset of receiver domains, all associated with phytochrome-like receptor kinases. Physicochemically, these 7 or 8 additional residues constitute a mostly hydrophobic patch, that interacts with the hydrophobic core of the dimer interface (see 4.3.2).

As described (4.3.2), the interaction between the related protomers is highly specific and not a result of the crystal packing. Furthermore, it was possible to deduce a conserved motive of amino acids in the primary structure that are involved in the molecular recognition between the monomers. This was accomplished by performing a protein BLAST search (<http://www.ncbi.nlm.nih.gov/BLAST>) with the sequence of RcpA and collecting those homologues that share most of the amino acids involved in dimerization. Figure 5.3 shows the alignment of most of the found receiver domains. Additionally, the genetic context of all RECs was revisited to see, what the associated [cotranscribed] receptors of those receivers were. The result was that all RECs, sharing the dimerization motive, either are co-transcribed with a bacterial phytochrome like receptor, a putative phytochrome like receptor or an unidentified receptor histidine kinase.

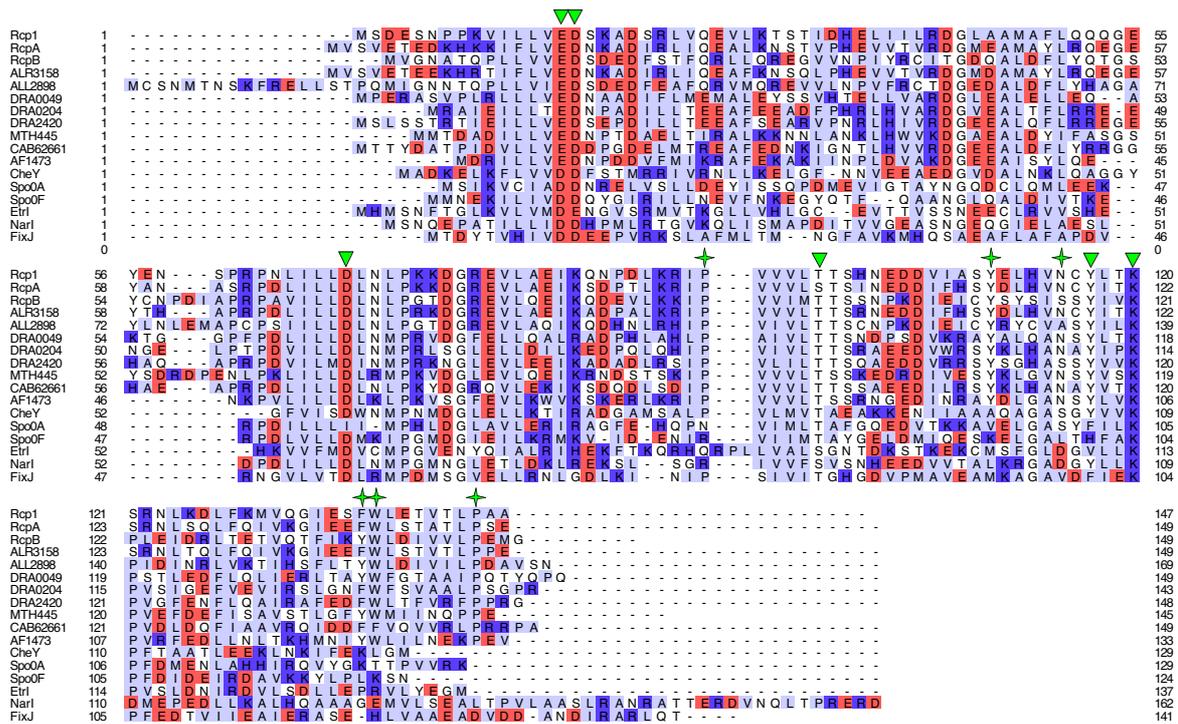


Figure. 5.3. **Homology alignment of RRs of the CheY type.** Highly conserved active site residues are marked with arrow heads. Green stars indicate residues, conserved in dimer formation in RcpA and B. Rcp1 (*Synechocystis sp.*, SLR0474), RcpA and B (*Calothrix sp.*), and RRs from *Anabaena sp.* [ALR3158, ALL2898] and *D. radiodurans* [DRA0049, DRA0204, DR2420] are highly homologous and also co-transcribed with putative phytochrome-homologous receptor histidine kinases. Several other RRs from *Methanobacterium thermoautotrophicum* [MTH445], *S. coelicolor* [CAB62661] or *Archeoglobus fulgidus*, AF1473] also share the dimerization motifs, whereas non of the structurally characterized CheY-homologues such as SpoOA [*B. stearrowthermophilus*], SpoOF [*B. subtilis*], Etr1 [*A. thaliana*], receiver domains of NarI [*E. coli*] and FixJ [*Sinorhizobium meliloti*] or CheY [*E. coli*] itself possess these conserved residues.

Among other receiver domains of the CheY type that are structurally characterized, three, eukaryotic Etr1 [Muller-Dieckmann, et al., 1999] and prokaryotic FixJN or N-SpoOA [Birck, et al., 1999, Gouet, et al., 1999, Lewis, et al., 1999], also seem to form homodimers, apparently, in a phosphorylation dependent manner, but the interaction of the respective protomers is essentially different from that found in RcpA and RcpB. For the unphosphorylated N-terminal receiver domain N-SpoOA of the sporulation response regulator SpoOA [*Bacillus stearrowthermophilus*], it was shown that dimer formation is mediated by the swapping of a helical domain, and that the dimeric species represents the active form of SpoOA that regulates transcription [Lewis, et al., 2002]. The receiver domain FixJN of the transcriptional activator FixJ from *Sinorhizobium meliloti* associates via the  $\beta$ 4/H5 regions, also in a phosphorylation dependent manner [Da Re, et al., 1999].

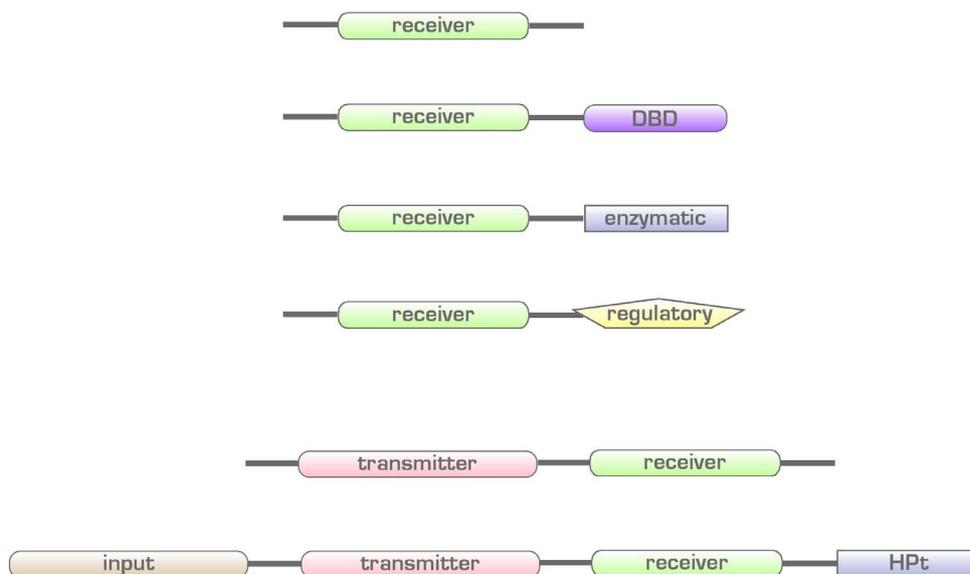


Figure 5.4. **Functional domains and modular organization in two-component signaling.**

Receiver domains appear in a multitude of combinations with other functional domains. In some cases, response regulators are single domain phosphorylatable receiver modules, such as RcpA, B and CheY. Usually they are combined with additional output domains such as DNA binding domains (DBD), enzymatic domains or regulatory domains that are regulated by the receiver domain. Some receiver modules are parts of multi-domain histidine kinases (transmitter) or receptor histidine kinases.

Unlike the above mentioned receiver modules N-SpoOA and FixJN and most other REC domains that are part of multi domain proteins (figure 5.4), RcpA and B are single, domain proteins, which possess no obvious output domain. It is presently unknown how the signal that is perceived by the phytochrome photoreceptor and transmitted to the receiver, is processed to downstream components and what these specifically are. As will be discussed in section 5.1.3, dimer formation might be triggered by phosphorylation, and dimers might represent the active signaling species of the Rcps.

The results from gelfiltration experiments with both RRs showed that under the given conditions the proteins were always dimeric with apparent molecular weights of twice the theoretical values [4.1.1, 4.1.2 and 4.2.1]. Thus, at first glance, the data did not promote dynamic dimerization in a phosphorylation dependent manner [ see also 5.1.3]. Nevertheless, interaction between the related protomers is highly specific. Comparison of the amino acid sequences shows, that RcpA shares a higher homology with Rcp1 from *Synechocystis sp.* PCC6803 than with its paralogue RcpB, and in parts, this difference between RcpA and RcpB certainly accounts for the specificity of the interaction between the cognate partners. Many of the non-conserved residues mediating the stabilization and formation of the dimer are similar in RcpA and Rcp1, but different in RcpB. A sequence comparison of RRs associated with (putative) bacterial

phytochromes and other, more CheY homologous RRs gives rise to the suggestion that the conserved C-terminal extensions in this family not only stabilize the formation of the dimeric protein, but also play a key role in the specific recognition among the protomers.

### 5.1.3 Activation of the cyanobacterial response regulators RcpA and RcpB

The central question for response regulator functions undoubtedly focuses on the structural changes accompanying the transition from inactive to active state, and their interaction with downstream components. Literature provides a vast amount of information concerning the mechanics involved in the switching event. Common to all is the conformational change that is undergone by a highly conserved threonine or serine residue in  $\beta$ -strand 4 (T87 in CheY) upon phosphorylation. But from this point on, the picture seems to blur in the different systems. For CheY, it is proposed that the threonine flip generates space for a conserved aromatic residue in  $\beta$ 5 which can then occupy a more buried rotatory conformation and thus enables downstream components (FliM, CheZ) to bind to the H4- $\beta$ 5-H5 surface [Lee, et al., 2001, Lee, et al., 2001, Zhu, et al., 1997]. On the other hand, the same conformational changes were not detected in the constitutively active mutant CheY<sub>D13K</sub> [Jiang, et al., 1997]. Another mutant, CheY<sub>T87A</sub>, can still be activated by phosphorylation, even though the key residue THR87 is replaced by a nonpolar alanine [Appleby and Bourret, 1998]. Receiver domains of other two-component systems, such as FixJN and N-SpoOA show very different effects upon phosphorylation. Whereas the receiver domain FixJN undergoes significant conformational changes like CheY and homodimerizes, to finally become part of a transcriptional activation complex [Birck, et al., 1999], the structure of N-SpoOA remains relatively unperturbed [Lewis, et al., 1999].

#### 5.1.3.1 The active and inactive state in RcpA and RcpB

The crystal structures of unphosphorylated RcpA and phosphorylated RcpB permit the comparison of the putative inactive state of RcpB, with the putative active state of RcpA, from a structural perspective. Due to the high sequence, structural and functional homology shared by the two Rcps, some characteristic features of this subclass of phytochrome-associated regulatory domains could be deduced.

Strikingly, the conformational differences in the active sites and associated structural elements in phospho-RcpA and apo-RcpB are only of minor nature, when compared to CheY, and less predominant. Figure 5.5 illustrates the conformations of active site residues and the rotatory position of the conserved tyrosine (Y119<sub>RcpA</sub> and Y118<sub>RcpB</sub>) of both Rcps, homologous to those residues involved in the allosteric control of the CheA

vs. FlIM binding affinity of CheY. In both Rcps, the respective tyrosine side chain occupies the inward directed position, similar to that in active CheY (Fig. 5.2.d).

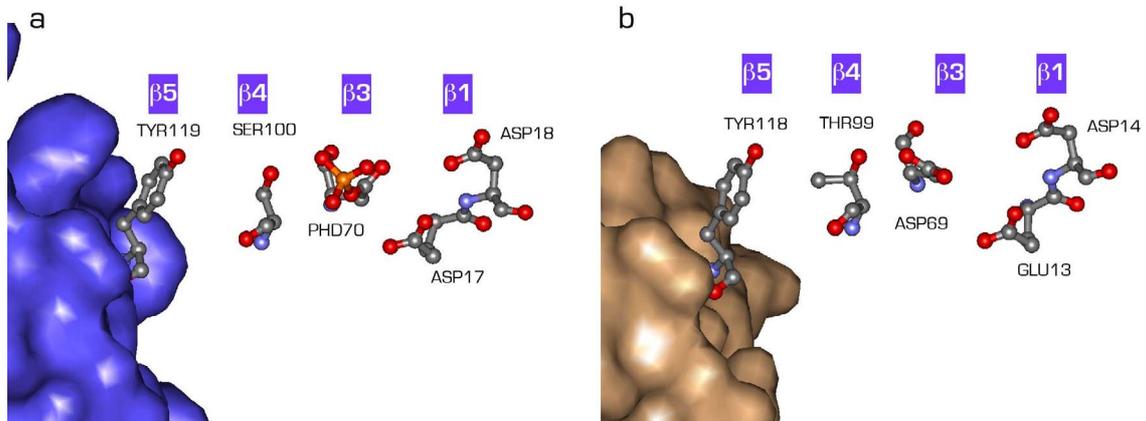


Figure 5.5. **Active state RcpA vs. inactive state RcpB.** The conserved key residues involved in activation are shown as ball and stick models for (a) phopho-RcpA and (b) apo-RcpB, together with part of the adjacent molecular surface of the second protomer in each dimer. In both structures, the tyrosine side chain adopts the inward rotated position, and SER100 and THR99 both form hydrogen bonds, with either the phosphate oxygen of PHD70 or the carboxylate oxygen of ASP69, respectively.

Major conformational differences between the apo- and the phospho-form predominantly include displacements in the region forming the acidic crevice at the surface as described in section 4.3.3. As referenced in the previous section, the phosphorylation of CheY involves the movement of two key residues, T87 and Y106, a major step in activation, which allosterically alters the characteristics of surface H4- $\beta$ 5-H5. In RcpA and B, the very same surface is involved in establishing dimer formation, and the respective key residues, T100 (RcpA) or S99 (RcpB) and Y119<sub>RcpA</sub> or Y118<sub>RcpB</sub> in both forms, occupy positions similar to those in activated CheY (chapter 4, figure 4.22 and 5.2). In the case of the cyanobacterial Rcps, dimerization prevents, for sterical reasons, that the corresponding solvent exposed rotamer (Fig. 5.2.b) can form in the inactive apo-RcpB.

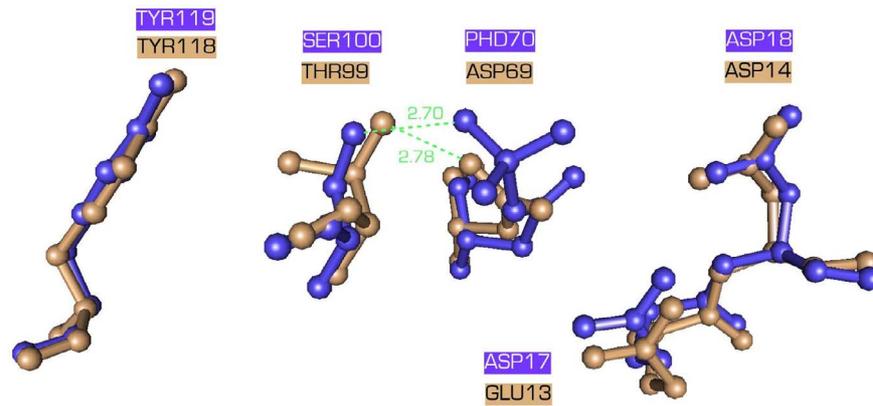


Figure 5.6. **Overlaid active site residues of phospho-RcpA and RcpB.** Ball and Stick representation of RcpA (blue) and RcpB (caramel) active site residues. Superimposing the active site residues of phospho-RcpA and apo-RcpB clearly demonstrates the degree to which the active site of RcpB has adopted the activated conformation in the dimer.

In phospho-RcpA, these findings are in perfect agreement with the postulated mechanism (5.1.1). The active form of RcpA has an inward rotated tyrosine side chain which enables a second activated RcpA molecule to bind against the H4- $\beta$ 5-H5 surface and to form a dimer, comparable to the binding of CheA by CheY. Applying the mechanism for activation of CheY and thinking one step ahead, this would also imply that dephosphorylation would, in turn, switch RcpA *off* and shift the equilibrium to the monomeric state, because the tyrosine side chain is no longer stabilized by a H-bond to the phosphate group and preferentially adopts the outward rotated conformation. Like in CheY, the monomeric form, with a bare H4- $\beta$ 5-H5 face and a solvent exposed tyrosine side chain, would offer an altered binding surface for the interaction with the receptor or downstream components. Taking into account the findings by Hübschmann et al. [Hübschmann, et al., 2001], who have shown that both Rcps are phosphorylated by their cognate photoreceptors CphA and CphB, a signaling model, as depicted in figure 5.7 can be summarized.

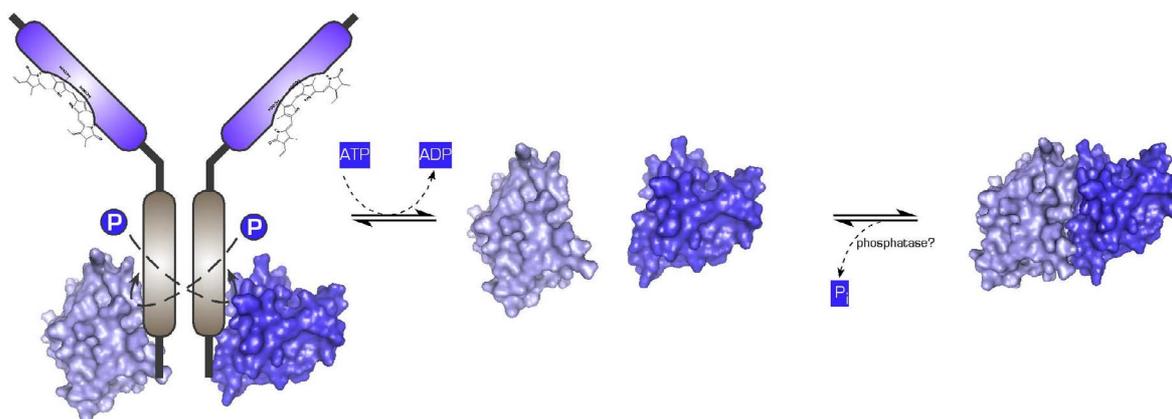


Figure 5.7. **Postulated mechanism for the functioning of phytochrome RRs.** Under *in vivo* conditions, the usual working flow of an Rcp involves formation and disintegration of a dimeric species in a phosphorylation dependent manner. The light activated dimeric receptor phytochrome autocatalytically phosphorylates a conserved histidine and binds the RR monomers to a dedicated, yet unknown site in the C-terminal domain. The following His-Asp phosphotransfer event generates the activated phospho-species of the receiver module, which dissociates from the receptor and forms a dimer with a second activated molecule. Dephosphorylation by either spontaneous hydrolysis or the action of a yet unknown phosphatase lowers the binding affinity of the receiver for its homodimeric partner and regenerates the monomeric RRs.

In contrast to that, in the crystal structure of dimeric apo-RcpB, the active site looks very similar to that of phospho-RcpA without the presence of a phosphate group as shown in figure 5.6. If dimerization is a dynamic, phosphorylation dependent process like postulated above, this would imply that in the case of RcpB, the dimeric state is non-native and presumably resulting from the experimental conditions during purification and/or crystallization. The experimental conditions for crystallization require high protein concentration to be maintained, which made up to 200 mM in the case of the Rcps. Results published by Hübschmann et al. concerning the natural abundance of Cph1 in *Synechocystis sp.* PCC6803, state that the *in vivo* concentration of the receptor might be as low as tens to hundreds of copies Cph1 per cell (Hübschmann et al., 2001, Eur. J. Biochem., 268). As the cognate response regulator Rcp1 is expressed under the control of the same *cph1* promotor, it is reasonable to assume a concentration for Rcp1 that lies in the same order of magnitude or is not very different from that. Apparently, in the case of CheY, the *in vivo* concentration in the cell was determined to be approx. 10  $\mu$ M (Bray and Bourret, 1995, Bray, et al., 1993, Gegner, et al., 1992) which is by far higher than that of Rcp1, but still  $10^4$  fold below the concentration, used in the crystallization setup. It is therefore not surprising, that the equilibrium of Rcp association might be shifted to the dimeric side. In the case of apo-RcpB, dimerization forces the active site residues to adopt a *pseudo* active conformation by a reverse allosteric mechanism. By forcing the tyrosine side chain inwards, the docking of two

monomers forces the contacting residues to also adopt a more suitable position and thereby imitating a phosphoryl-bound active site conformation. Association of two monomeric RcpB molecules apparently occurs under non-native conditions like high concentrations and without the need for phosphorylation.

### 5.1.3.2 Dynamic or static?

The above made assumptions for a dynamic dimerization mechanism are based on X-ray and gelfiltration data and on the activation mechanism of CheY. Due to the lack of detailed biochemical information on the system, the postulated mechanism for Rcp-switching only constitutes a possible hypothesis, albeit one that is certainly more favorable. At a swift glance, the X-ray structures of both Rcps and data from size exclusion chromatography [4.1.1, 4.1.2 and 4.2.1] definitely promote the alternative assumption that dimers are the constitutive and stable form of the Rcps *in vivo*. Thus, the central question to the Rcp-signaling mechanism that arises is, whether dimer formation allosterically influences the conformation in the active site and alters its catalytic activity, or whether the active site conformation, dependent on the phosphorylation state, affects dimer formation. A possible speculation would be the following. As deduced from the comparison of the active sites of RcpA and RcpB, the conformation of inactive RcpB resembles more the active state of RcpA. The most obvious mechanism would of course be the induction of a conformational change in the protein upon phosphorylation, which shifts the equilibrium to the active and dimeric state. Although the crystallographic data does not support this, the finding that both forms, active and inactive, appear as dimers in the crystals, might as well be an experimental artifact, as outlined above. On the other hand, an alternative interpretation of the data would be that Rcps are constitutively dimeric *in vivo*, irrespective of phosphorylation. Whether this is of any biological relevance, can not be resolved at the present state, but one is definitely invited to speculate on the benefit that constitutive dimerization might mean for the system. The formation of the dimeric species could have a reverse effect on the active site conformation and thereby alter the phosphoryl transfer activity of the regulator. In the case of the transcription factor OmpR for example, a two domain protein with an amino-terminal CheY-like receiver domain and a carboxy-terminal DNA-binding domain, the following has been described. Primarily, phosphorylation of OmpR increases the affinity of the DNA-binding domain for the *ompF* and *ompC* regulatory sites [Aiba, et al., 1989, Aiba, et al., 1989]. But, as found by Kenny and colleagues [Ames, et al., 1999], the reverse holds true as well and DNA-binding of OmpR can stimulate OmpR phosphorylation by its cognate kinase EnvZ or the small molecule phosphodonor acetyl phosphate. As highlighted by Buckler et al. [Buckler, et al., 2000], dimerization of the receiver domains of OmpR upon binding to their target DNA site could have an intermolecular effect and allosterically influence catalysis at the active

site. The effect of this allosteric activation of the phosphoryl transfer activity would definitely influence the level of phosphorylation of the intracellular pool of the respective receiver domain. As to whether this has any relevance for the phytochrome associated response regulators, and whether dimer formation allosterically alters their phosphoryl transfer activity definitely, remains to be elucidated through future experimental work.

### 5.1.3.3 Conclusion

A mechanism for the signaling of bacterial phytochrome response regulators has been postulated on the basis of X-ray and chromatography data. Furthermore, the results indicate clearly that the specificity of dimer formation is driven by a highly conserved complementarity of the interface region, generating a tightly interacting homodimer. The interfaces in dimeric RcpA and B bury a surface area of 1000 Å<sup>2</sup> and 1175 Å<sup>2</sup>, respectively, which is in the same range as observed in the complexes of CheY and FlhM (1110 Å<sup>2</sup>), CheY and CheA (P2 domain) (1200 Å<sup>2</sup>), but larger than dimeric FixJ (880 Å<sup>2</sup>) [Da Re, et al., 1999, Lee, et al., 2001, Welch, et al., 1998]. If the dimeric form is constitutive however, as promoted by a strict interpretation of the data, it remains to be investigated what the characteristic features of the activated species of these receiver domains might be and how the signal is processed to further downstream partners. So far, unfortunately, there is no biochemical data available on RcpA/B downstream interacting components that could relate the introduced system to one, that is better understood.

## 5.2 59 and 65 kD N-terminal fragments of oat phytochrome A

### 5.2.1 Expression and purification

Both N-terminal receptor domains of oat PhyA were expressed in recombinant yeasts and purified to homogeneity by immobilized metal ion affinity (IMAC) and size exclusion chromatography. In most cases, apo proteins were reassembled with the chromophore PCB in a crude cell lysate, prior to purification. In few cases, the assembly was performed with highly purified apo-proteins in a Tris HCl buffered saline solution. As the formation rate of holo-protein was comparable to those performed in crude cellular extracts, this demonstrated that the autocatalytic activity was fully established in the purified apo-protein. It is thus generally possible to purify recombinant phytochromes as the apo form, which extremely simplifies the operative effort that usually accompanies the purification of the photosensitive holo-proteins. The affinity purification step could be optimized, by applying an imidazole gradient for elution of bound molecules to such an extent, that the eluted proteins were essentially pure (> 95 %) from other protein impurities as judged from SDS-PAGE after one single step. The final gel filtration step further improved the quality of the samples by removing residual impurities and yielding protein preparations that were considered highly pure as assessed from SDS-PAGE analysis and the spectroscopic data. The SAR of the final samples were in the range of 1.5 to 1.6 for both fragments. These values could not further be improved by additional purification methods such as ion exchange chromatography and should therefore represent the maximum SAR, achievable for these recombinant constructs. Unfortunately, the apparently pure samples revealed an immanent inhomogeneity when analyzed by isoelectric focussing techniques (see 4.4.1.1). Both fragments were separated in polyacrylamide gels containing ampholytes (pH range 3-7) and unveiled multiple coomassie stained bands, distributed over nearly 3 pH values. This has to be ascribed to severe microheterogeneities in the proteins. These differences in charge can be the result of a multitude of processes. Common sources for microheterogeneity are misfolding or partial unfolding, conformational variation in the tertiary structure, specific or unspecific oligomerization or the presence of flexible domains. Furthermore, incomplete specific or unspecific post-translational modifications, partial binding of ligands or other small molecules (phosphorylation, covalent binding of protease inhibitors), fragmentation and chemical alteration, like partial oxidation or deamidation, are possible sources for inhomogeneity, which are usually difficult to identify, and, more important, to prevent. In the case of the two phy fragments, most probably a combination of several processes has to be considered. The gel filtration indicated a homogenous size distribution in the sample and anion exchange chromatography also supported the presence of one major population in the sample (data not shown). Both methods are not resolutive enough to discriminate minor differences in charge, size or

folding states. The results obtained from ESI-MS [4.4.1.1] promoted the presence of multiple species in the samples, but showed no indication for one major and clearly defined molecular mass, probably as a result from terminal and unspecific proteolysis. The fact that the apparent molecular weights of the phy fragments, as deduced from gel filtration experiments [4.4.1.1] reflected not the theoretically expected values, but were between the monomeric and dimeric masses (81 kD for 59 kD PhyA and 110 kD for 65 kD PhyA), originated from either a dynamic equilibrium between monomeric and multimeric species or conformationally flexible (partially unfolded) molecules which dynamically transform between structurally close conformers. If and to what extent the molecules with different pI, resolved in the IEF experiment, derived from either chemically different species or/and differently folded and partially unfolded forms, that exhibit altered surface electrostatic potentials could not be adjudged. The results from NMR experiments that further support these findings will be discussed in chapter 5.2.3.

### 5.2.2 Photoactivity of phy fragments

The autocatalytic activity to incorporate and covalently attach the offered bilin chromophore was obviously unaffected by the inhomogeneity of the proteins. The holo-protein usually formed in the range of minutes and was spectrally undistinguishable from native phytochromes, irrespective of a slight hypsochromic shift in the absorption maxima, due to the use of PCB instead of the native P $\Phi$ B. The holo-proteins of both fragments also were photochemically active and fully photoreversible. This implied that the protein fragments apparently adopt a native fold at least in domains, which are responsible for chromophore incorporation and photoreversibility. By performing circular dichroic spectroscopy, it could be unequivocally shown that the 59 and 65 kD fragments possess secondary structure elements (see 4.4.2), and that the protein backbone is affected by the isomerization of the chromophore scaffold upon light absorption. The 65 kD PhyA domain seems to undergo a reversible refolding that is triggered by the light absorption event. By comparing the CD signals for 65 kD PhyA with those obtained from 59 kD PhyA, it can be hypothesized that mainly the N-terminal 65 amino acids contribute to the difference in the 65 kD P<sub>r</sub> and P<sub>fr</sub> signals. The 59 kD PhyA, lacking this peptide, shows a comparably weak alteration of the CD spectrum upon light switching from P<sub>r</sub> to P<sub>fr</sub> form. Secondary structure analysis ascribe with high confidence a predominantly helical secondary fold to the N-terminal peptide. It has previously been stated that a refolding event in this N-terminal helix accompanies the Pr to Pfr conversion ([Chen, 1997 #167; Farrens, 1992 #164]. The N-terminus of plant PhyA contains two functional domains (see 2.2.1), one necessary for conformational stability and biological activity (residues 13-62), and the other involved in attenuating phytochrome responses (residues 6-12) as described by Jordan et al. (Jordan, et al., 1996). In fact, quite recent studies (Fankhauser, 2002, Sweere, et al., 2001) have

attested a specific interaction with the response regulator *ARR4* from *Arabidopsis*, which binds to the amino-terminus of PhyB and inhibits dark reversion of the  $P_{fr}$  form to inactive  $P_r$ .

### 5.2.3 Crystallization trials and NMR-spectroscopy

As outlined in 4.4.4, the 59 and 65 kD PhyA fragments were extensively tested for their suitability for crystallization. Initial screenings with sparse matrix approaches as described in the experimental section, led to unordered precipitation in all cases where supersaturation was achieved in the droplets. None of the crystallization setups produced crystalline-like arrangements or microcrystals. This behavior of the protein must be attributed to the insufficient quality of the preparations in terms of conformational, chemical or physico-chemical homogeneity. As described in the previous paragraph, the samples had an inherent heterogeneity that could not be reduced by modified purification protocols, such as separating differently charged subspecies by additional anion exchange chromatography, extended use of protease inhibitors, reduced temperatures and dark light conditions during purification, applying different cell lysis methods (glass mill, ultra-turrax homogenizer) and the addition of stabilizing agents like 2-mercaptoethanol, electrolytes or glycerol. The concept of purity in macromolecular crystallography is different to that, usually claimed in other fields of experimental biochemistry. Proteins that appear to be pure as judged from SDS gel electrophoresis and possess high enzymatic activity are often sufficient for many experimental applications. The questions addressed in most experimental setups other than crystallogenesis do not require extensive physicochemical analysis of the macromolecules involved or exhaustive analysis is neglected because the vital conditions (e.g. purity > 90 %, enzymatic activity) are fulfilled. In the case of the two PhyA fragments, the spectroscopic characteristics and the autocatalytic activity for chromophore incorporation definitely constitute adequate means for judging the functional integrity of recombinant probes. Nevertheless, crystal growth demands a higher grade of homogeneity, and above all, a homogeneous population of conformations.

These prerequisites can be excellently monitored by nuclear magnetic resonance spectroscopy which is a powerful technique for investigating the conformational homogeneity and the folding state of a given macromolecule. Other than circular dichroic methods which give an average image of the overall secondary structure, NMR has the ability to also monitor the tertiary fold of a molecule and can reveal dynamic phenomena, besides giving information on impurities in greater detail. From a simple  $^1\text{H}$  NMR experiment, the specific and stable folding of a protein can be deduced by analyzing the chemical shift regions of the amide/aromatic and methyl groups. A pure, well-folded globular protein shows a distinct chemical shift dispersion and well defined,

sharp NMR signals, although the latter also depends on the size of the protein and the magnetic field strength. The broadening of peaks and a narrow distribution (clustering), as was found in experiments performed with 59 kD PhyA [see below], is highly indicative for conformational flexibility in the protein backbone or aggregation [Hill and DeGrado, 2000]. Fig. 5.8 shows examples of one-dimensional  $^1\text{H}$  NMR spectra of a folded and of an unfolded protein. The  $^1\text{H}$  NMR spectrum recorded from the PhyA fragment is shown in Fig. 4.31.

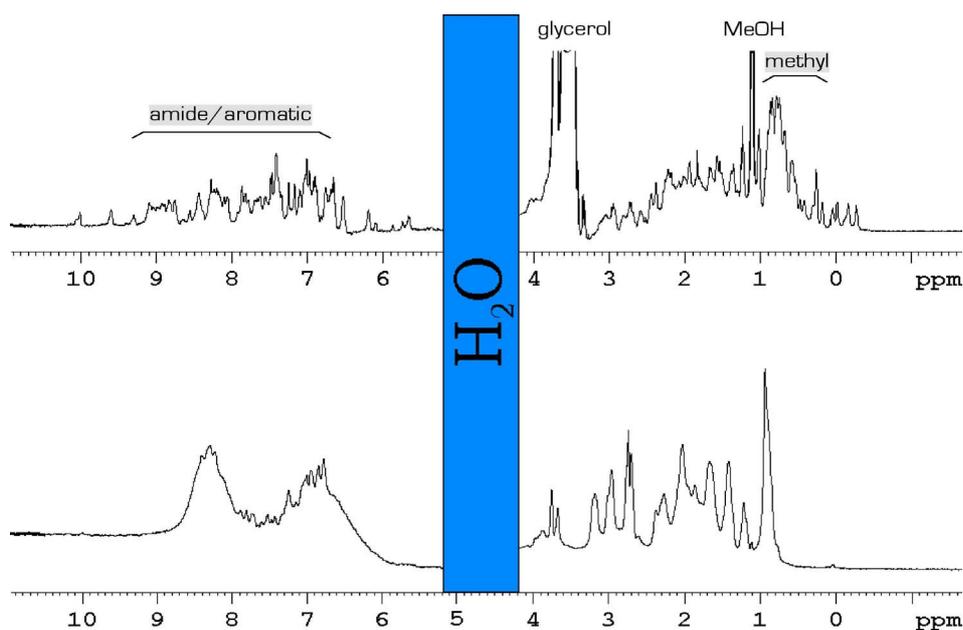


Figure 5.8. **Typical  $^1\text{H}$ -D NMR experiments** of a folded (barnase inhibitor barstar, top) and an unfolded (human p27, bottom) protein. The denoted regions for amide/aromatic and methyl chemical shifts are indicative for the folding state of a protein sample and exhibit an obvious signal dispersion in the folded state. Spectra are courtesy of the NMR Service Group of the Max-Planck-Institute for Biochemistry, Martinsried.

The marked regions for the amide/aromatic and methyl chemical shifts in both phy fragments indicate that the 3D-structural "composition" [conformational population] of the sample is not homogeneous. The sample either represent a heterogeneous manifold with structurally related conformers, probably as a result of unfolded, partially folded and misfolded molecules, and/or a specifically unfolded state, generally referred to as "molten globule" that exhibits great conformational flexibility in solution, the dynamics of which lie beyond the time window of the NMR technique. In order to determine whether folding is in some way related to the activation state of the receptor domains, and to gain a more detailed view of the folding state of the protein samples, a two dimensional  $^{15}\text{N}$  heteronuclear NMR experiment [ $^1\text{H}$ - $^{15}\text{N}$  HSQC-NMR] was performed. As outlined in 3.5.3, this NMR experiment allows to resolve the amide region

of a protein in an additional dimension according to the  $^{15}\text{N}$  chemical shifts. Every amino acid in a given protein produces one signal, and the chemical shift depends on the local magnetic environment for each amide group. Therefore, a specifically folded protein with residues at dedicated, more or less restrained positions, shows a pattern of well distributed peaks in the  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum, whereas those of unfolded proteins cluster in few regions. Figure 5.9 shows typical  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra for folded and unfolded proteins.

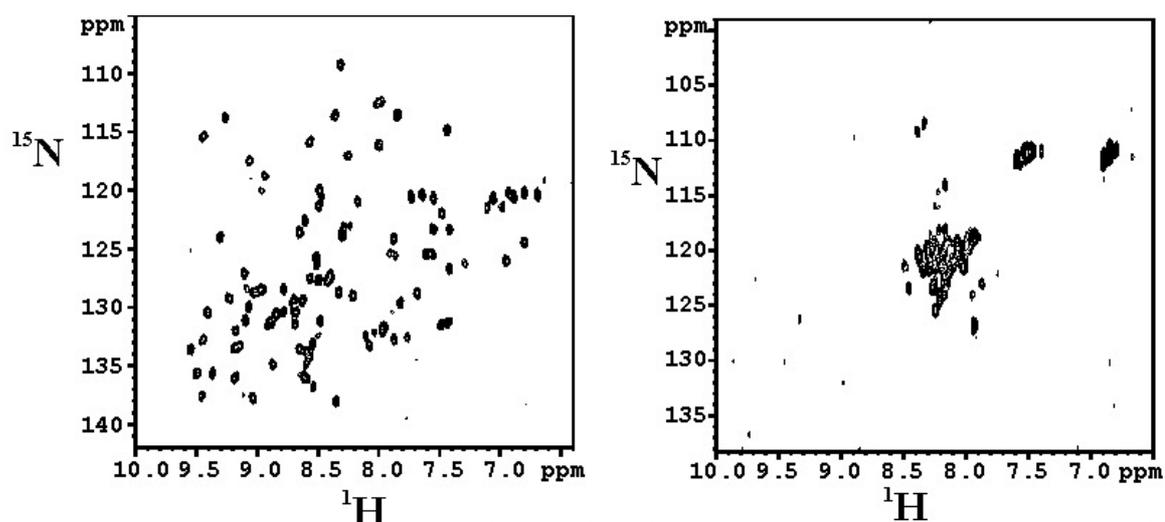


Figure 5.9.  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra. Shown are typical examples of two dimensional HSQC NMR spectra of (a) a folded protein [actin binding protein ABP-120], with signals that are well-spread over a broad range, and (b) of an unfolded protein [human retinoblastoma protein], with clustered signals and little distribution. Spectra are courtesy of the NMR Service Group of the Max-Planck-Institute for Biochemistry, Martinsried.

The 59 kD PhyA was completely labeled with the nitrogen isotope  $^{15}\text{N}$  as described and  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectra were recorded on a 750 MHz NMR spectrometer for both, the  $P_r$  and the  $P_{fr}$  form. Both spectra revealed, what was already assumed from the one-dimensional experiment, namely the characteristic NMR features for an unfolded or highly dynamic protein. The spectra for the  $P_r$  and  $P_{fr}$  form show only minor differences in chemical shifts, and in both forms, the amide signals cluster and overlap in the region around 8.0 - 8.5 ppm, typical for unfolded proteins. Additionally, the side chain amide signals of glutamine and asparagine residues are clustered around 6.8 ppm and 7.5 ppm, also characteristic for random coil proteins (Wüthrich, 1986). It has been described that proteins exist in a conformationally flexible native state, while being fully functional. In the case of photoactive yellow protein (PYP) for example, it was shown by HSQC and other NMR experiments, that the pB intermediate of the PYP photocycle exhibits a large degree of disorder due to the presence of multiple conformers, that

exchange on a millisecond time scale [Dux, et al., 1998, Lee, et al., 2001, Rubinstenn, et al., 1998]. In other cases, it has been described that minor changes in the protein (point mutations) or its surrounding (pH shifts) can produce molten globule-like states [Li, et al., 1999, Matthews, et al., 2000, Yuan, et al., 1999, Zhang and Forman-Kay, 1997]. This is not surprising, as specific protein folding and secondary and tertiary structure stability are cooperative characteristics, and minor perturbations, like the missing of a single amino acid side chain that normally would be involved in H-bond formation, can induce partial unfolding of domains [Privalov, 1979]. Altogether, the experimental results from IEF, crystallization and NMR studies constitute more than persuasive arguments that demonstrate the inappropriateness of the 59 kD fragment of oat PhyA for crystallization, produced under the described conditions for expression and purification. Whether heterogeneity forms during purification or is already present in the intact cells is not known, but it was clearly shown that crystallization is not achievable with these constructs. As mentioned, the stability of the three dimensional structure of a protein is a cooperative effect and the sum of multiple parameters that add up to stabilize a certain structural fold with minimized energy. The fragments investigated during this study were designed according to results from unspecific proteolysis experiments with trypsin protease, and thus do not necessarily represent stable, compact subdomains. The absence of any domain or amino acid, normally involved in stabilizing the correct fold, might result in a destabilizing effect and lead to misfolded or unfolded molecules. On the other hand, the stable and specific folding of the native receptor might involve the presence of other components of the signaling pathway like Arr4 [Fankhauser, 2002, Sweere, et al., 2001], PIF3 [Zhu, et al., 2000] or NDPK1 [Quail, 2000], see 2.1.4, and the formation of complexes with the latter might be necessary for a rigid ensemble, that translocates to the nucleus and integrates into a transcriptional regulatory complex.

An alternative approach for determining the structure of plant phytochromes by X-ray crystallography that was also commenced in work, was the cloning and expression of full length *A. thaliana* PhyA and PhyB phytochromes and their common signaling partner PIF3 [Ni, et al., 1998]. These should be tested for their ability to form specific, stable complexes and that might be more suitable for crystallization. Additionally, these full length receptors could be subjected to limited proteolysis experiments with unspecific proteases, other than trypsin (subtilisin, thermolysin, papain etc.) in order to produce and identify more compact, stable sub-domains. Unfortunately, completion of these experiments was beyond the time limit of this work.

## 6 Bibliography

**Adler, A. J., Greenfield, N. J. and Fasman, G. D.** (1973) Circular dichroism and optical rotatory dispersion of proteins and polypeptides. *Methods Enzymol* **27**, 675-735.

**Aiba, H., Nakasai, F., Mizushima, S. and Mizuno, T.** (1989) Evidence for the physiological importance of the phosphotransfer between the two regulatory components, EnvZ and OmpR, in osmoregulation in *Escherichia coli*. *J Biol Chem* **264**, 14090-4.

**Aiba, H., Nakasai, F., Mizushima, S. and Mizuno, T.** (1989) Phosphorylation of a bacterial activator protein, OmpR, by a protein kinase, EnvZ, results in stimulation of its DNA-binding ability. *J Biochem (Tokyo)* **106**, 5-7.

**Ames, S. K., Frankema, N. and Kenney, L. J.** (1999) C-terminal DNA binding stimulates N-terminal phosphorylation of the outer membrane protein regulator OmpR from *Escherichia coli*. *Proc Natl Acad Sci USA* **96**, 11792-7.

**Appleby, J. L. and Bourret, R. B.** (1998) Proposed signal transduction role for conserved CheY residue Thr87, a member of the response regulator active-site quintet. *J Bacteriol* **180**, 3563-9.

**Arndt, U. W. and Wonacott, A.** (1978) The rotation method in crystallography. *North-Holland Publishers, Amsterdam*.

**Bazzi, M. D. and Woody, R. W.** (1985) Oriented secondary structure in integral membrane proteins. I. Circular dichroism and infrared spectroscopy of cytochrome oxidase in multilamellar films. *Biophys J* **48**, 957-66.

**Birck, C., Mourey, L., Gouet, P., Fabry, B., Schumacher, J., Rousseau, P., Kahn, D. and Samama, J. P.** (1999) Conformational changes induced by phosphorylation of the FixJ receiver domain. *Structure Fold Des* **7**, 1505-15.

**Blair, D. F.** (1995) How bacteria sense and swim. *Annu Rev Microbiol* **49**, 489-522.

**Blundell, T. L. and Johnson, L. N.** *Protein Crystallography* [Academic Press, New York, 1976].

**Bodenhausen, G. a. R., D. J.** (1980) Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69**, 185-189.

**Bray, D. and Bourret, R. B.** (1995) Computer analysis of the binding reactions leading to a transmembrane receptor-linked multiprotein complex involved in bacterial chemotaxis. *Mol Biol Cell* **6**, 1367-80.

**Bray, D., Bourret, R. B. and Simon, M. I.** (1993) Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Mol Biol Cell* **4**, 469-82.

**Bren, A. and Eisenbach, M.** (1998) The N terminus of the flagellar switch protein, FliM, is the binding domain for the chemotactic response regulator, CheY. *J Mol Biol* **278**, 507-14.

**Brünger, A. T.** (1992) The Free R value: a Novel Statistical Quantity for Assessing the Accuracy of Crystal Structures. *Nature* **355**, 472-5.

**Brünger, A. T.** (1993) Assessment of phase accuracy by cross validation: The free R value. Methods and Applications. *Acta Crystallogr D Biol Crystallogr* **49**, 24-36.

**Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. and Warren, G. L.** (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-21.

**Brünger, A. T., Adams, P. D. and Rice, L. M.** (1997) New applications of simulated annealing in X-ray crystallography and solution NMR. *Structure* **5**, 325-36.

**Buckler, D. R., Anand, G. S. and Stock, A. M.** (2000) Response-regulator phosphorylation and activation: a two-way street? *Trends Microbiol* **8**, 153-6.

**Budisa, N., Steipe, B., Demange, P., Eckerskorn, C., Kellermann, J. and Huber, R.** (1995) High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur J Biochem* **230**, 788-96.

**Buttler, W. L., Norris, K. H., Siegelman, S. B. and Hendricks, S. B.** (1959) Detection, assay and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc. Natl. Acad. Sci. USA* **45**, 1703-1708.

**Cashmore, A. R., Jarillo, J. A., Wu, Y. J. and Liu, D.** (1999) Cryptochromes: blue light receptors for plants and animals. *Science* **284**, 760-5.

**Chen, E., Lapko, V. N., Song, P. S. and Kliger, D. S.** (1997) Dynamics of the N-terminal alpha-helix unfolding in the photoreversion reaction of phytochrome A. *Biochemistry* **36**, 4903-8.

**Chory, J.** (1997) Light modulation of vegetative development. *Plant Cell* **9**, 1225-34.

**Chory, J., Chatterjee, M., Cook, R. K., Elich, T., Fankhauser, C., Li, J., Nagpal, P., Neff, M., Pepper, A., Poole, D., Reed, J. and Vitart, V.** (1996) From seed germination to flowering, light controls plant development via the pigment phytochrome. *Proc Natl Acad Sci U S A* **93**, 12066-71.

**Collaborative Computational Project, N.** (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst. D50*, 760-763.

**Da Re, S., Schumacher, J., Rousseau, P., Fourment, J., Ebel, C. and Kahn, D.** (1999) Phosphorylation-induced dimerization of the FixJ receiver domain. *Mol Microbiol* **34**, 504-11.

**Deleage, G. and Geourjon, C.** (1993) An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. *Comput Appl Biosci* **9**, 197-9.

**Devlin, P. F., Patel, S. R. and Whitelam, G. C.** (1998) Phytochrome E influences internode elongation and flowering time in Arabidopsis. *Plant Cell* **10**, 1479-87.

**Devlin, P. F., Robson, P. R., Patel, S. R., Goosey, L., Sharrock, R. A. and Whitelam, G. C.** (1999) Phytochrome D acts in the shade-avoidance syndrome in Arabidopsis by controlling elongation growth and flowering time. *Plant Physiol* **119**, 909-15.

**Drenth, J.** *Principles of Protein X-ray Crystallography* (Springer, New York, 1999).

**Dutta, R., Qin, L. and Inouye, M.** (1999) Histidine kinases: diversity of domain organization. *Mol Microbiol* **34**, 633-40.

**Dux, P., Rubinstenn, G., Vuister, G. W., Boelens, R., Mulder, F. A., Hard, K., Hoff, W. D., Kroon, A. R., Crielaard, W., Hellingwerf, K. J. and Kaptein, R.** (1998) Solution structure and backbone dynamics of the photoactive yellow protein. *Biochemistry* **37**, 12689-99.

**Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A. and Danielson, M. A.** (1997) The two-component signaling pathway of bacterial chemotaxis: a molecular view of signal transduction by receptors, kinases, and adaptation enzymes. *Annu Rev Cell Dev Biol* **13**, 457-512.

**Fankhauser, C.** (2002) Light perception in plants: cytokinins and red light join forces to keep phytochrome B active. *Trends Plant Sci* **7**, 143-5.

**Fankhauser, C., Yeh, K. C., Lagarias, J. C., Zhang, H., Elich, T. D. and Chory, J.** (1999) PKS1, a substrate phosphorylated by phytochrome that modulates light signaling in Arabidopsis. *Science* **284**, 1539-41.

**Foussard, M., Cabantous, S., Pedelacq, J., Guillet, V., Tranier, S., Mourey, L., Birck, C. and Samama, J.** (2001) The molecular puzzle of two-component signaling cascades. *Microbes Infect* **3**, 417-24.

**French, S. W., K.** (1978) On the treatment of negative intensity observations. *Acta Cryst.* **A34**, 517-525.

**Ganguli, S., Wang, H., Matsumura, P. and Volz, K.** (1995) Uncoupled phosphorylation and activation in bacterial chemotaxis. The 2.1 Å structure of a threonine to isoleucine mutant at position 87 of CheY. *J Biol Chem* **270**, 17386-93.

**Gärtner, W., Hill, C., Worm, K., Braslavsky, S. E. and Schaffner, K.** (1996) Influence of expression system on chromophore binding and preservation of spectral properties in recombinant phytochrome A. *Eur J Biochem* **236**, 978-83.

**Gegner, J. A., Graham, D. R., Roth, A. F. and Dahlquist, F. W.** (1992) Assembly of an MCP receptor, CheW, and kinase CheA complex in the bacterial chemotaxis signal transduction pathway. *Cell* **70**, 975-82.

**Giacovazzo, C., Monaco, H. L. and Viterbo, D.** *Fundamentals of Crystallography* (Oxford Univ. Press, Oxford, 1992).

**Gouet, P., Fabry, B., Guillet, V., Birck, C., Mourey, L., Kahn, D. and Samama, J. P.** (1999) Structural transitions in the FixJ receiver domain. *Structure Fold Des* **7**, 1517-26.

**Grebe, T. W. and Stock, J. B.** (1999) The histidine protein kinase superfamily. *Adv Microb Physiol* **41**, 139-227.

**Green, R. M. and Tobin, E. M.** (1999) Loss of the circadian clock-associated protein 1 in Arabidopsis results in altered clock-regulated gene expression. *Proc Natl Acad Sci U S A* **96**, 4176-9.

**Guex, N. and Peitsch, M. C.** (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-23.

**Hahn, T.** *International Tables for Crystallography* (ed. T. Hahn) [Kluwer Academic Publishers, Dordrecht, 1995].

**Hellingwerf, K. J., Hoff, W. D. and Crielgaard, W.** (1996) Photobiology of microorganisms: how photosensors catch a photon to initialize signalling. *Mol Microbiol* **21**, 683-93.

**Hendrickson, W. A., Horton, J. R. and LeMaster, D. M.** (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *Embo J* **9**, 1665-72.

**Hendrickson, W. A., Smith, J. L. and Sheriff, S.** (1985) Direct phase determination based on anomalous scattering. *Methods Enzymol* **115**, 41-55.

**Hill, C., Gärtner, W., Towner, P., Braslavsky, S. E. and Schaffner, K.** (1994) Expression of phytochrome apoprotein from *Avena sativa* in *Escherichia coli* and formation of photoactive chromoproteins by assembly with phycocyanobilin. *Eur J Biochem* **223**, 69-77.

**Hill, R. B. and DeGrado, W. F.** (2000) A polar, solvent-exposed residue can be essential for native protein structure. *Structure Fold Des* **8**, 471-9.

**Hoppe, W.** (1957) Die Faltmolekülmethode und ihre Anwendung in der röntgenographischen Konstitutionsanalyse von Bifforin (C<sub>10</sub>H<sub>20</sub>O<sub>4</sub>). *Z. Elektrochemie* **61**, 1076-1083.

**Hübschmann, T., Jorissen, H. J., Börner, T., Gärtner, W. and Tandeau de Marsac, N.** (2001) Phosphorylation of proteins in the light-dependent signalling pathway of a filamentous cyanobacterium. *Eur J Biochem* **268**, 3383-9.

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

**Jiang, M., Bourret, R. B., Simon, M. I. and Volz, K.** (1997) Uncoupled phosphorylation and activation in bacterial chemotaxis. The 2.3 Å structure of an aspartate to lysine mutant at position 13 of CheY. *J Biol Chem* **272**, 11850-5.

**Johnson, W. C., Jr.** (1985) Circular dichroism and its empirical application to biopolymers. *Methods Biochem Anal* **31**, 61-163.

**Johnson, W. C., Jr.** (1988) Secondary structure of proteins through circular dichroism spectroscopy. *Annu Rev Biophys Biophys Chem* **17**, 145-66.

**Johnson, W. C., Jr.** (1990) Protein secondary structure and circular dichroism: a practical guide. *Proteins* **7**, 205-14.

**Jones, D. T.** (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202.

**Jones, T. A., Zou, J.Y., Cowan, S.W. and Kjeldgaard, M.** (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst A* **47**, 110-119.

**Jordan, E. T., Cherry, J. R., Walker, J. M. and Vierstra, R. D.** (1996) The amino-terminus of phytochrome A contains two distinct functional domains. *Plant J* **9**, 243-57.

**Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S.** (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res* **3**, 109-36.

**Karle, J.** (1980) Some developments in anomalous dispersion for the structural investigation of macromolecular systems in biology. *International Journal of Quantum Chemistry: Quantum Biology Symposium* **7**, 357-367.

**Kendrick, R. E. and Kronenberg, G. H. M.** *Photomorphogenesis in plants* (Kluwer Dordrecht, 1994).

**Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K., Schäfer, E. and Nagy, F.** (1999) Light quality-dependent nuclear import of the plant photoreceptors phytochrome A and B. *Plant Cell* **11**, 1445-56.

**Kleiner, O., Kircher, S., Harter, K. and Batschauer, A.** (1999) Nuclear localization of the Arabidopsis blue light receptor cryptochrome 2. *Plant J* **19**, 289-96.

**Kleywegt, G. J. and Brunger, A. T.** (1996) Checking your imagination: applications of the free R value. *Structure* **4**, 897-904.

**Kleywegt, G. J. and Jones, T. A.** (1998) Databases in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **54**, 1119-31.

**Kleywegt, G. J. and Read, R. J.** (1997) Not your average density. *Structure* **5**, 1557-69.

**Kneip, C., Hildebrandt, P., Schlamann, W., Braslavsky, S. E., Mark, F. and Schaffner, K.** (1999) Protonation state and structural changes of the tetrapyrrole chromophore during the Pr → Pfr phototransformation of phytochrome: a resonance Raman spectroscopic study. *Biochemistry* **38**, 15185-92.

**Krah, M., Marwan, W., Vermeglio, A. and Oesterhelt, D.** (1994) Phototaxis of *Halobacterium salinarium* requires a signalling complex of sensory rhodopsin I and its methyl-accepting transducer HtrI. *Embo J* **13**, 2150-5.

**Kraulis, P. J.** (1991) MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures. *Journal of Applied Crystallography* **24**, 946-950.

**Kufer, W. and Scheer, H.** (1979a) Chemical modification of biliprotein chromophores. *Z Naturforsch [C]* **34**, 776-81.

**Kufer, W. and Scheer, H.** (1979b) Studies on plant bile pigments, VII. Preparation and characterization of phycobiliproteins with chromophores chemically modified by reduction. *Hoppe Seylers Z Physiol Chem* **360**, 935-56.

**Laemmli, U. K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.

**Lagarias, J. C. and Lagarias, D. M.** (1989) Self-assembly of synthetic phytochrome holoprotein in vitro. *Proc Natl Acad Sci U S A* **86**, 5778-80.

**Larsen, T. A., Olson, A. J. and Goodsell, D. S.** (1998) Morphology of protein-protein interfaces. *Structure Fold Des* **6**, 421-7.

**Lawrence, M. C. and Colman, P. M.** (1993) Shape complementarity at protein/protein interfaces. *J Mol Biol* **234**, 946-50.

**Lee, B. C., Croonquist, P. A., Sosnick, T. R. and Hoff, W. D.** (2001) PAS domain receptor photoactive yellow protein is converted to a molten globule state upon activation. *J Biol Chem* **276**, 20821-3.

**Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Berry, E. A. and Wemmer, D. E.** (2001) Crystal structure of activated CheY. Comparison with other activated receiver domains. *J Biol Chem* **276**, 16425-31.

**Lee, S. Y., Cho, H. S., Pelton, J. G., Yan, D., Henderson, R. K., King, D. S., Huang, L., Kustu, S., Berry, E. A. and Wemmer, D. E.** (2001) Crystal structure of an activated response regulator bound to its target. *Nat Struct Biol* **8**, 52-6.

**Leslie, A. G. W.** (1992) Int. CCP4/ESF-EACMB. *Newslett. Protein Crystallogr.* **26**.

**Lewis, R. J., Brannigan, J. A., Muchova, K., Barak, I. and Wilkinson, A. J.** (1999) Phosphorylated aspartate in the structure of a response regulator protein. *J Mol Biol* **294**, 9-15.

**Lewis, R. J., Scott, D. J., Brannigan, J. A., Ladds, J. C., Cervin, M. A., Spiegelman, G. B., Hoggett, J. G., Barak, I. and Wilkinson, A. J.** (2002) Dimer formation and transcription activation in the sporulation response regulator SpoOA. *J Mol Biol* **316**, 235-45.

**Li, D., Allen, D. L., Harvey, S., Perrino, F. W., Schaaper, R. M. and London, R. E.** (1999) A preliminary CD and NMR study of the Escherichia coli DNA polymerase III theta subunit. *Proteins* **36**, 111-6.

**Lin, C.** (2002) Phototropin blue light receptors and light-induced movement responses in plants. *Science STKE* **2002**, E5.

**Liu, Y., Levit, M., Lurz, R., Surette, M. G. and Stock, J. B.** (1997) Receptor-mediated protein kinase activation and the mechanism of transmembrane signaling in bacterial chemotaxis. *Embo J* **16**, 7231-40.

**Martinez-Garcia, J. F., Huq, E. and Quail, P. H.** (2000) Direct targeting of light signals to a promoter element-bound transcription factor. *Science* **288**, 859-63.

**Marwan, W., Bibikov, S. I., Montrone, M. and Oesterhelt, D.** (1995) Mechanism of photosensory adaptation in Halobacterium salinarium. *J Mol Biol* **246**, 493-9.

**Matthews, B. W.** (1968) Solvent content of protein crystals. *J Mol Biol* **33**, 491-7.

**Matthews, J. M., Norton, R. S., Hammacher, A. and Simpson, R. J.** (2000) The single mutation Phe173 → Ala induces a molten globule-like state in murine interleukin-6. *Biochemistry* **39**, 1942-50.

**McCleary, W. R. and Stock, J. B.** (1994) Acetyl phosphate and the activation of two-component response regulators. *J Biol Chem* **269**, 31567-72.

**McDonell, M. W., Simon, M. N. and Studier, F. W.** (1977) Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J Mol Biol* **110**, 119-46.

**McEvoy, M. M., Bren, A., Eisenbach, M. and Dahlquist, F. W.** (1999) Identification of the binding interfaces on CheY for two of its targets, the phosphatase CheZ and the flagellar switch protein fliM. *J Mol Biol* **289**, 1423-33.

**McMichael, R. W., Jr. and Lagarias, J. C.** (1990) Phosphopeptide mapping of Avena phytochrome phosphorylated by protein kinases in vitro. *Biochemistry* **29**, 3872-8.

**McPhearson, A.** *Crystallization of Biological Macromolecules* (John Wiley & Sons, New York, 1999).

**Millar, A. J., McGrath, R. B. and Chua, N. H.** (1994) Phytochrome phototransduction pathways. *Annu Rev Genet* **28**, 325-49.

- Morrison, T. B. and Parkinson, J. S.** (1994) Liberation of an interaction domain from the phosphotransfer region of CheA, a signaling kinase of *Escherichia coli*. *Proc Natl Acad Sci U S A* **91**, 5485-9.
- Mozley, D., Remberg, A. and Gärtner, W.** (1997) Large-scale generation of affinity-purified recombinant phytochrome chromopeptide. *Photochem Photobiol* **66**, 710-5.
- Muller-Dieckmann, H. J., Grantz, A. A. and Kim, S. H.** (1999) The structure of the signal receiver domain of the *Arabidopsis thaliana* ethylene receptor ETR1. *Structure Fold Des* **7**, 1547-56.
- Nagy, F. and Schäfer, E.** (2000a) Control of nuclear import and phytochromes. *Curr Opin Plant Biol* **3**, 450-4.
- Nagy, F. and Schäfer, E.** (2000b) Nuclear and cytosolic events of light-induced, phytochrome-regulated signaling in higher plants. *Embo J* **19**, 157-63.
- Navaza, J.** (1994) AMoRe: an automated package for molecular replacement. *Acta Cryst. A* **50**, 157-163.
- Ni, M., Tepperman, J. M. and Quail, P. H.** (1998) PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. *Cell* **95**, 657-67.
- Ni, M., Tepperman, J. M. and Quail, P. H.** (1999) Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature* **400**, 781-4.
- Nicholls, A., Sharp, K. and Honig, B.** (1991) *PROTEINS* **11**, 281-296.
- Otwinowski, Z. and Minor, W.** in *Methods in Enzymology, Macromolecular Crystallography, part A* (eds. C. W. Carter and R. M. Sweet) 307-326 [Academic Press, 1997].
- Parkinson, J. S. and Blair, D. F.** (1993) Does *E. coli* have a nose? *Science* **259**, 1701-2.
- Parkinson, J. S. and Kofoed, E. C.** (1992) Communication modules in bacterial signaling proteins. *Annu Rev Genet* **26**, 71-112.
- Patterson, A. L.** (1934) A Fourier series method for the determination of the components of interatomic distances in crystals. *Phys. Rev.* **46**, 372-376.
- Perrakis, A., Morris, R. and Lamzin, V. S.** (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**, 458-63.
- Perutz, M. F.** (1956) isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Cryst.* **9**, 867-873.
- Privalov, P. L.** (1979) Stability of proteins: small globular proteins. *Adv Protein Chem* **33**, 167-241.
- Quail, P. H.** (1991) Phytochrome: a light-activated molecular switch that regulates plant gene expression. *Annu Rev Genet* **25**, 389-409.

**Quail, P. H.** (1994) Photosensory perception and signal transduction in plants. *Curr Opin Genet Dev* **4**, 652-61.

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

**Quail, P. H.** (2000) Phytochrome-interacting factors. *Semin Cell Dev Biol* **11**, 457-66.

**Quail, P. H.** (2002) Photosensory perception and signalling in plant cells: new paradigms? *Curr Opin Cell Biol* **14**, 180-8.

**Quail, P. H., Boylan, M. T., Parks, B. M., Short, T. W., Xu, Y. and Wagner, D.** (1995) Phytochromes: photosensory perception and signal transduction. *Science* **268**, 675-80.

**Romanos, M. A., Scorer, C. A. and Clare, J. J.** (1992) Foreign gene expression in yeast: a review. *Yeast* **8**, 423-88.

**Rossmann, M. G.** *The Molecular Replacement Method* (Gordon and Breach, New York, 1972).

**Rossmann, M. G. a. B., D.M.** (1962) The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst.* **15**, 24-31.

**Rubinstenn, G., Vuister, G. W., Mulder, F. A., Dux, P. E., Boelens, R., Hellingwerf, K. J. and Kaptein, R.** (1998) Structural and dynamic changes of photoactive yellow protein during its photocycle in solution. *Nat Struct Biol* **5**, 568-70.

**Sakamoto, K. and Nagatani, A.** (1996) Nuclear localization activity of phytochrome B. *Plant J* **10**, 859-68.

**Sambrook, J., Fritsch, E. F. and Maniatis, T.** in *Molecular Cloning A.2* (, Cold Spring Harbor, 1989).

**Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. and Roe, B. A.** (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J Mol Biol* **143**, 161-78.

**Sanger, F., Nicklen, S. and Coulson, A. R.** (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* **74**, 5463-7.

**Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., Carre, I. A. and Coupland, G.** (1998) The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **93**, 1219-29.

**Smith, H.** (2000) Phytochromes and light signal perception by plants - an emerging synthesis. *Nature* **407**, 585-91.

**Sousa, R.** (1995) Use of glycerol, polyols and other protein structure stabilizing agents in protein chrySTALLIZATION. *Acta Crystallogr D Biol Crystallogr* **51**, 271-277.

- Southern, E.** (1979) Gel electrophoresis of restriction fragments. *Methods Enzymol* **68**, 152-76.
- Sreerama, N. W., R. W.** (1993) A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem* **209**, 32-44.
- Stewart, R. C.** (1997) Kinetic characterization of phosphotransfer between CheA and CheY in the bacterial chemotaxis signal transduction pathway. *Biochemistry* **36**, 2030-40.
- Stock, A. M., Martinez-Hackert, E., Rasmussen, B. F., West, A. H., Stock, J. B., Ringe, D. and Petsko, G. A.** (1993) Structure of the Mg<sup>2+</sup>-bound form of CheY and mechanism of phosphoryl transfer in bacterial chemotaxis. *Biochemistry* **32**, 13375-80.
- Stock, A. M., Mottonen, J. M., Stock, J. B. and Schutt, C. E.** (1989) Three-dimensional structure of CheY, the response regulator of bacterial chemotaxis. *Nature* **337**, 745-9.
- Stock, A. M., Robinson, V. L. and Goudreau, P. N.** (2000) Two-component signal transduction. *Annu Rev Biochem* **69**, 183-215.
- Stock, J. and Da Re, S.** (2000) Signal transduction: response regulators on and off. *Curr Biol* **10**, R420-4.
- Stout, G. H. and Jensen, L. H.** *X-Ray Structure Determination* (Wiley & Sons, New York, 1989).
- Studier, F. W., Rosenberg, A. H., Dunn, J. J. and Dubendorff, J. W.** (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**, 60-89.
- Sweere, U., Eichenberg, K., Lohrmann, J., Mira-Rodado, V., Baurle, I., Kudla, J., Nagy, F., Schäfer, E. and Harter, K.** (2001) Interaction of the response regulator ARR4 with phytochrome B in modulating red light signaling. *Science* **294**, 1108-11.
- Vierstra, R. D. and Davis, S. J.** (2000) Bacteriophytochromes: new tools for understanding phytochrome signal transduction. *Semin Cell Dev Biol* **11**, 511-21.
- Wang, Z. Y., Kenigsbuch, D., Sun, L., Harel, E., Ong, M. S. and Tobin, E. M.** (1997) A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* **9**, 491-507.
- Wang, Z. Y. and Tobin, E. M.** (1998) Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**, 1207-17.
- Welch, M., Chinardet, N., Mourey, L., Birck, C. and Samama, J. P.** (1998) Structure of the CheY-binding domain of histidine kinase CheA in complex with CheY. *Nat Struct Biol* **5**, 25-9.
- Welch, M., Oosawa, K., Aizawa, S. I. and Eisenbach, M.** (1994) Effects of phosphorylation, Mg<sup>2+</sup>, and conformation of the chemotaxis protein CheY on its binding to the flagellar switch protein FliM. *Biochemistry* **33**, 10470-6.

**Whitelam, G. C., Patel, S. and Devlin, P. F.** (1998) Phytochromes and photomorphogenesis in Arabidopsis. *Philos Trans R Soc Lond B Biol Sci* **353**, 1445-53.

**Wilson, A. J. C.** (1949) The probability distribution of X-ray intensities. *Acta Cryst.* **2**, 318-321.

**Wong, Y. S., Cheng, H. C., Walsh, D. A. and Lagarias, J. C.** (1986) Phosphorylation of Avena phytochrome in vitro as a probe of light-induced conformational changes. *J Biol Chem* **261**, 12089-97.

**Wüthrich, K.** *NMR of Proteins and Nucleic Acids* (John Wiley, 1986).

**Yamaguchi, R., Nakamura, M., Mochizuki, N., Kay, S. A. and Nagatani, A.** (1999) Light-dependent translocation of a phytochrome B-GFP fusion protein to the nucleus in transgenic Arabidopsis. *J Cell Biol* **145**, 437-45.

**Yang, J. T., Wu, C. S. and Martinez, H. M.** (1986) Calculation of protein conformation from circular dichroism. *Methods Enzymol* **130**, 208-69.

**Yeh, K. C. and Lagarias, J. C.** (1998) Eukaryotic phytochromes: light-regulated serine/threonine protein kinases with histidine kinase ancestry. *Proc Natl Acad Sci U S A* **95**, 13976-81.

**Yeh, K. C., Wu, S. H., Murphy, J. T. and Lagarias, J. C.** (1997) A cyanobacterial phytochrome two-component light sensory system. *Science* **277**, 1505-8.

**Yuan, C., Byeon, I. J., Poi, M. J. and Tsai, M. D.** (1999) Structural analysis of phospholipase A2 from functional perspective. 2. Characterization of a molten globule-like state induced by site-specific mutagenesis. *Biochemistry* **38**, 2919-29.

**Zhang, O. and Forman-Kay, J. D.** (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry* **36**, 3959-70.

**Zhu, X., Amsler, C. D., Volz, K. and Matsumura, P.** (1996) Tyrosine 106 of CheY plays an important role in chemotaxis signal transduction in Escherichia coli. *J Bacteriol* **178**, 4208-15.

**Zhu, X., Rebello, J., Matsumura, P. and Volz, K.** (1997) Crystal structures of CheY mutants Y106W and T87I/Y106W. CheY activation correlates with movement of residue 106. *J Biol Chem* **272**, 5000-6.

**Zhu, X., Volz, K. and Matsumura, P.** (1997) The CheZ-binding surface of CheY overlaps the CheA- and FliM-binding surfaces. *J Biol Chem* **272**, 23758-64.

**Zhu, Y., Tepperman, J. M., Fairchild, C. D. and Quail, P. H.** (2000) Phytochrome B binds with greater apparent affinity than phytochrome A to the basic helix-loop-helix factor PIF3 in a reaction requiring the PAS domain of PIF3. *Proc Natl Acad Sci U S A* **97**, 13419-24.

## 7 Appendix

## Sparse Matrix Screening

## Hampton Screen I

1. 0.02 M Calcium Chloride dihydrate	1. 0.1 M Sodium Acetate trihydrate pH 4.6	1. 30% v/v 2-Methyl-2,4-pentanediol
2. None	2. None	2. 0.4 M Potassium Sodium Tartrate tetrahydrate
3. None	3. None	3. 0.4 M mono-Ammonium dihydrogen Phosphate
4. None	4. 0.1 M Tris Hydrochloride pH 8.5	4. 2.0 M Ammonium Sulfate
5. 0.2 M tri-Sodium Citrate dihydrate	5. 0.1 M HEPES - Na pH 7.5	5. 30% v/v 2-Methyl-2,4-pentanediol
6. 0.2 M Magnesium Chloride hexahydrate	6. 0.1 M Tris Hydrochloride pH 8.5	6. 30% w/v Polyethylene Glycol 4000
7. None	7. 0.1 M Sodium Cacodylate pH 6.5	7. 1.4 M Sodium Acetate trihydrate
8. 0.2 M tri-Sodium Citrate dihydrate	8. 0.1 M Sodium Cacodylate pH 6.5	8. 30% v/v iso-Propanol
9. 0.2 M Ammonium Acetate	9. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	9. 30% w/v Polyethylene Glycol 4000
10. 0.2 M Ammonium Acetate	10. 0.1 M Sodium Acetate trihydrate pH 4.6	10. 30% w/v Polyethylene Glycol 4000
11. None	11. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	11. 1.0 M mono-Ammonium dihydrogen Phosphate
12. 0.2 M Magnesium Chloride hexahydrate	12. 0.1 M HEPES - Na pH 7.5	12. 30% v/v iso-Propanol
13. 0.2 M tri-Sodium Citrate dihydrate	13. 0.1 M Tris Hydrochloride pH 8.5	13. 30% v/v Polyethylene Glycol 400
14. 0.2 M Calcium Chloride dihydrate	14. 0.1 M HEPES - Na pH 7.5	14. 28% v/v Polyethylene Glycol 400
15. 0.2 M Ammonium Sulfate	15. 0.1 M Sodium Cacodylate pH 6.5	15. 30% w/v Polyethylene Glycol 8000
16. None	16. 0.1 M HEPES - Na pH 7.5	16. 1.5 M Lithium Sulfate monohydrate
17. 0.2 M Lithium Sulfate monohydrate	17. 0.1 M Tris Hydrochloride pH 8.5	17. 30% Polyethylene Glycol 4000
18. 0.2 M Magnesium Acetate tetrahydrate	18. 0.1 M Sodium Cacodylate pH 6.5	18. 20% Polyethylene Glycol 8000
19. 0.2 M Ammonium Acetate	19. 0.1 M Tris Hydrochloride pH 8.5	19. 30% v/v iso-Propanol
20. 0.2 M Ammonium Sulfate	20. 0.1 M Sodium Acetate trihydrate pH 4.6	20. 25% w/v Polyethylene Glycol 4000
21. 0.2 M Magnesium Acetate tetrahydrate	21. 0.1 M Sodium Cacodylate pH 6.5	21. 30% v/v 2-Methyl-2,4-pentanediol
22. 0.2 M Sodium Acetate trihydrate	22. 0.1 M Tris Hydrochloride pH 8.5	22. 30% w/v Polyethylene Glycol 4000
23. 0.2 M Magnesium chloride hexahydrate	23. 0.1 M HEPES - Na pH 7.5	23. 30% v/v Polyethylene Glycol 400
24. 0.2 M Calcium Chloride dihydrate	24. 0.1 M Sodium Acetate trihydrate pH 4.6	24. 20% v/v iso-Propanol
25. None	25. 0.1 M Imidazole pH 6.5	25. 1.0 M Sodium Acetate trihydrate
26. 0.2 M Ammonium Acetate	26. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	26. 30 % v/v 2-Methyl-2,4-pentanediol
27. 0.2 M tri-Sodium Citrate dihydrate	27. 0.1 M HEPES - Na pH 7.5	27. 20% v/v iso-Propanol
28. 0.2 M Sodium Acetate trihydrate	28. 0.1 M Sodium Cacodylate pH 6.5	28. 30% w/v Polyethylene Glycol 8000
29. None	29. 0.1 M HEPES - Na pH 7.5	29. 0.8 M Potassium Sodium Tartrate tetrahydrate
30. 0.2 M Ammonium Sulfate	30. None	30. 30% w/v Polyethylene Glycol 8000
31. 0.2 M Ammonium Sulfate	31. None	31. 30% w/v Polyethylene Glycol 4000
32. None	32. None	32. 2.0 M Ammonium Sulfate
33. None	33. None	33. 4.0 M Sodium Formate
34. None	34. 0.1 M Sodium Acetate trihydrate pH 4.6	34. 2.0 M Sodium Formate
35. None	35. 0.1 M HEPES - Na pH 7.5	35. 0.8 M mono-Sodium dihydrogen phosphate 0.8 M mono-Potassiumdihydrogen phosphate
36. None	36. 0.1 M Tris Hydrochloride pH 8.5	36. 8% w/v Polyethylene Glycol 8000
37. None	37. 0.1 M Sodium Acetate trihydrate pH 4.6	37. 8% w/v Polyethylene Glycol 4000
38. None	38. 0.1 M HEPES - Na pH 7.5	38. 1.4 M tri-Sodium Citrate dihydrate
39. None	39. 0.1 M HEPES - Na pH 7.5	39. 2% v/v Polyethylene Glycol 400, 2.0 M Ammonium Sulfate
40. None	40. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	40. 20% v/v iso-Propanol, 20% w/v Polyethylene Glycol 4000
41. None	41. 0.1 M HEPES - Na pH 7.5	41. 10% v/v iso-Propanol, 20% w/v Polyethylene Glycol 4000
42. 0.05 M mono-Potassium dihydrogen Phosphate	42. None	42. 20% w/v Polyethylene Glycol 8000
43. None	43. None	43. 30% w/v Polyethylene Glycol 1500
44. None	44. None	44. 0.2 M Magnesium Formate
45. 0.2 M Zinc Acetate dihydrate	45. 0.1 M Sodium Cacodylate pH 6.5	45. 18% w/v Polyethylene Glycol 8000
46. 0.2 M Calcium Acetate hydrate	46. 0.1 M Sodium Cacodylate pH 6.5	46. 18% w/v Polyethylene Glycol 8000
47. None	47. 0.1 M Sodium Acetate trihydrate pH 4.6	47. 2.0 M Ammonium Sulfate
48. None	48. 0.1 M Tris Hydrochloride pH 8.5	48. 2.0 M mono-Ammonium dihydrogen Phosphate
49. 1.0 M Lithium Sulfate monohydrate	49. None	49. 2% w/v Polyethylene Glycol 8000
50. 0.5 M Lithium Sulfate monohydrate	50. None	50. 15% w/v Polyethylene Glycol 8000

## Hampton Screen II

1. 2.0 M Sodium chloride	1. None	1. 10% w/v PEG 6000
2. 0.01 M Hexadecyltrimethylammonium Bromide	2. None	2 0.5 M Sodium Chloride, 0.01 M Magnesium Chloride hexahydrate
3. None	3. None	3. 25% v/v Ethylene Glycol
4. None	4. None	4. 35% v/v Dioxane
5. 2.0 M Ammonium Sulfate	5. None	5. 5% v/v iso-Propanol
6. None	6. None	6. 1.0 M Imidazole pH 7.0
7. None	7. None	7. 10% w/v Polyethylene Glycol 1000 10% w/v Polyethylene Glycol 8000
8. 1.5 M Sodium Chloride	8. None	8. 10% v/v Ethanol
9. None	9. 0.1 M Sodium Acetate trihydrate pH 4.6	9. 2.0 M Sodium Chloride
10. 0.2 M Sodium Chloride	10. 0.1 M Sodium Acetate trihydrate pH 4.6	10. 30% v/v MPD
11. 0.01 M Cobaltous Chloride hexahydrate	11. 0.1 M Sodium Acetate trihydrate pH 4.6	11. 1.0 M 1,6 Hexanediol
12. 0.1 M Cadmium Chloride dihydrate	12. 0.1 M Sodium Acetate trihydrate pH 4.6	12. 30% v/v Polyethylene Glycol 400
13. 0.2 M Ammonium Sulfate	13. 0.1 M Sodium Acetate trihydrate pH 4.6	13. 30% w/v Polyethylene Glycol Monomethyl Ether 2000
14. 0.2 M Potassium Sodium Tartrate tetrahydrate	14. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	14. 2.0 M Ammonium Sulfate
15. 0.5 M Ammonium Sulfate	15. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	15. 1.0 M Lithium Sulfate monohydrate
16. 0.5 M Sodium Chloride	16. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	16. 4% w/v Ethylene Imine Polymer
17. None	17. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	17. 35% v/v tert-Butanol
18. 0.01 M Ferric Chloride hexahydrate	18. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	18. 10% v/v Jeffamine M-600
19. None	19. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	19. 2.5 M 1,6 Hexanediol
20. None	20. 0.1 M MES pH 6.5	20. 1.6 M Magnesium Sulfate heptahydrate
21. 0.1 M Sodium dihydrogen phosphate mono 0.1 M mono-Potassium dihydrogen Phosphate	21. 0.1 M MES pH 6.5	21. 2.0 M Sodium Chloride
22. None	22. 0.1 M MES pH 6.5	22. 12% w/v Polyethylene Glycol 20,000
23. 1.6 M Ammonium Sulfate	23. 0.1 M MES pH 6.5	23. 10% v/v Dioxane
24. 0.05 M Cesium Chloride	24. 0.1 M MES pH 6.5	24. 30% v/v Jeffamine M-600
25. 0.01 M Cobaltous Chloride hexahydrate	25. 0.1 M MES pH 6.5	25. 1.8 M Ammonium Sulfate
26. 0.2 M Ammonium Sulfate	26. 0.1 M MES pH 6.5	26. 30% w/v Polyethylene Glycol Monomethyl Ether 5000
27. 0.01 M Zinc Sulfate heptahydrate	27. 0.1 M MES pH 6.5	27. 25% v/v Polyethylene Glycol Monomethyl Ether 550
28. None	28. None	28. 1.6 M tri-Sodium Citrate dihydrate pH 6.5
29. 0.5 M Ammonium Sulfate	29. 0.1 M HEPES pH 7.5	29. 30% v/v MPD
30. None	30. 0.1 M HEPES pH 7.5	30. 10% w/v Polyethylene Glycol 6000, 5% v/v MPD
31. None	31. 0.1 M HEPES pH 7.5	31. 20% v/v Jeffamine M-600
32. 0.1 M Sodium Chloride	32. 0.1 M HEPES pH 7.5	32. 1.6 M Ammonium Sulfate
33. None	33. 0.1 M HEPES pH 7.5	33. 2.0 M Ammonium Formate
34. 0.05 M Cadmium Sulfate hydrate	34. 0.1 M HEPES pH 7.5	34. 1.0 M Sodium acetate
35. None	35. 0.1 M HEPES pH 7.5	35. 70% v/v MPD
36. None	36. 0.1 M HEPES pH 7.5	36. 4.3 M Sodium Chloride
37. None	37. 0.1 M HEPES pH 7.5	37. 10% w/v Polyethylene Glycol 8000, 8% v/v Ethylene Glycol
38. None	38. 0.1 M HEPES pH 7.5	38. 20% w/v Polyethylene Glycol 10,000
39. 0.2 M Magnesium Chloride hexahydrate	39. 0.1 M TRIS pH 8.5	39. 3.4 M 1,6 Hexanediol
40. None	40. 0.1 M TRIS pH 8.5	40. 25% v/v tert-Butanol
41. 0.01 M Nickel(II) Chloride hexahydrate	41. 0.1 M TRIS pH 8.5	41. 1.0 M Lithium Sulfate monohydrate
42. 1.5 M Ammonium Sulfate	42. 0.1 M TRIS pH 8.5	42. 12% v/v Glycerol anhydrous
43. 0.2 M mono Ammonium dihydrogen Phosphate	43. 0.1 M TRIS pH 8.5	43. 50% v/v MPD
44. None	44. 0.1 M TRIS pH 8.5	44. 20% v/v Ethanol
45. 0.01 M Nickel(II) Chloride hexahydrate	45. 0.1 M TRIS pH 8.5	45. 20% w/v Polyethylene Glycol Monomethyl Ether 2000
46. 0.1 M Sodium Chloride	46. 0.1 M Bicine pH 9.0	46. 30% w/v Polyethylene Glycol Monomethyl Ether 550
47. None	47. 0.1 M Bicine pH 9.0	47. 2.0 M Magnesium Chloride hexahydrate
48. 2% v/v Dioxane	48. 0.1 M Bicine pH 9.0	48. 10% w/v Polyethylene Glycol 20,000

## Hampton Cryo Screen

1. 0.02 M Calcium Chloride dihydrate	1. 0.1 M Sodium Acetate trihydrate pH 4.6	1. 30% v/v 2-Methyl-2,4-pentanediol	1. None
2. None	2. None	2. 0.26 M K-Na-Tartrate tetrahydrate	2. 35% v/v
3. None	3. None	3. 0.26 M Ammonium dihydrogen Phosphate	3. 35% v/v
4. None	4. 0.075 M Tris Hydrochloride pH 8.5	4. 1.5 M Ammonium Sulfate	4. 25% v/v
5. 0.2 M tri-Sodium Citrate dihydrate	5. 0.1 M HEPES - Na pH 7.5	5. 30% v/v 2-Methyl-2,4-pentanediol	5. None
6. 0.16 M Mg Chloride hexahydrate	6. 0.08 M Tris Hydrochloride pH 8.5	6. 24% w/v Polyethylene Glycol 4000	6. 20% v/v
7. None	7. 0.07 M Sodium Cacodylate pH 6.5	7. 0.98 M Sodium Acetate trihydrate	7. 30% v/v
8. 0.14 M tri-Sodium Citrate dihydrate	8. 0.07 M Sodium Cacodylate pH 6.5	8. 21% v/v iso-Propanol	8. 30% v/v
9. 0.17 M Ammonium Acetate	9. 0.085 M tri-Sodium Citrate dihydrate pH 5.6	9. 25.5% w/v Polyethylene Glycol 4000	9. 15% v/v
10. 0.17 M Ammonium Acetate	10. 0.085 M Sodium Acetate trihydrate pH 4.6	10. 25.5% w/v Polyethylene Glycol 4000	10. 15% v/v
11. None	11. 0.07 M tri-Sodium Citrate dihydrate pH 5.6	11. 0.7 M Ammonium dihydrogen Phosphate	11. 30% v/v
12. 0.18 M Mg Chloride hexahydrate	12. 0.09 M HEPES - Na pH 7.5	12. 27% v/v iso-Propanol	12. 10% v/v
13. 0.2 M tri-Sodium Citrate dihydrate	13. 0.1 M Tris Hydrochloride pH 8.5	13. 30% v/v Polyethylene Glycol 400	13. None
14. 0.19 M Calcium Chloride dihydrate	14. 0.095 M HEPES - Na pH 7.5	14. 26.6% v/v Polyethylene Glycol 400	14. 5% v/v
15. 0.17 M Ammonium Sulfate	15. 0.085 M Sodium Cacodylate pH 6.5	15. 25.5% w/v Polyethylene Glycol 8000	15. 15% v/v
16. None	16. 0.075 M HEPES - Na pH 7.5	16. 1.125 M Lithium Sulfate monohydrate	16. 25% v/v
17. 0.17 M Lithium Sulfate monohydrate	17. 0.085 M Tris Hydrochloride pH 8.5	17. 25.5% Polyethylene Glycol 4000	17. 15% v/v
18. 0.16 M Mg Acetate tetrahydrate	18. 0.08 M Sodium Cacodylate pH 6.5	18. 16% Polyethylene Glycol 8000	18. 20% v/v
19. 0.16 M Ammonium Acetate	19. 0.08 M Tris Hydrochloride pH 8.5	19. 24% v/v iso-Propanol	19. 20% v/v
20. 0.16 M Ammonium Sulfate	20. 0.08 M Sodium Acetate trihydrate pH 4.6	20. 20% w/v Polyethylene Glycol 4000	20. 20% v/v
21. 0.2 M Mg Acetate tetrahydrate	21. 0.1 M Sodium Cacodylate pH 6.5	21. 30% v/v 2-Methyl-2,4-pentanediol	21. None
22. 0.17 M Sodium Acetate trihydrate	22. 0.085 M Tris Hydrochloride pH 8.5	22. 25.5% w/v Polyethylene Glycol 4000	22. 15% v/v
23. 0.2 M Mg Chloride hexahydrate	23. 0.1 M HEPES - Na pH 7.5	23. 30% v/v Polyethylene Glycol 400	23. None
24. 0.14 M Calcium Chloride dihydrate	24. 0.07 M Sodium Acetate trihydrate pH 4.6	24. 14% v/v iso-Propanol	24. 30% v/v
25. None	25. 0.07 M Imidazole pH 6.5	25. 0.7 M Sodium Acetate trihydrate	25. 30% v/v
26. 0.2 M Ammonium Acetate	26. 0.1 M tri-Sodium Citrate dihydrate pH 5.6	26. 30% v/v 2-Methyl-2,4-pentanediol	26. None
27. 0.14 M tri-Sodium Citrate dihydrate	27. 0.07 M HEPES - Na pH 7.5	27. 14% v/v iso-Propanol	27. 30% v/v
28. 0.17 M Sodium Acetate trihydrate	28. 0.085 M Sodium Cacodylate pH 6.5	28. 25.5% w/v Polyethylene Glycol 8000	28. 15% v/v
29. None	29. 0.065 M HEPES - Na pH 7.5	29. 0.52 M Potassium Na-Tartrate tetrahydrate	29. 35% v/v
30. 0.17 M Ammonium Sulfate	30. None	30. 25.5% w/v Polyethylene Glycol 8000	30. 15% v/v
31. 0.17 M Ammonium Sulfate	31. None	31. 25.5% w/v Polyethylene Glycol 4000	31. 15% v/v
32. None	32. None	32. 1.5 M Ammonium Sulfate	32. 25% v/v
33. None	33. None	33. 3.6 M Sodium Formate	33. 10% v/v
34. None	34. 0.07 M Sodium Acetate trihydrate pH 4.6	34. 1.4 M Sodium Formate	34. 30% v/v
35. None	35. 0.075 M HEPES - Na pH 7.5	35. 0.6 M NaH <sub>2</sub> PO <sub>4</sub> , 0.6 M KH <sub>2</sub> PO <sub>4</sub>	35. 25% v/v
36. None	36. 0.065 M Tris Hydrochloride pH 8.5	36. 5.2% w/v Polyethylene Glycol 8000	36. 35% v/v
37. None	37. 0.07 M Sodium Acetate trihydrate pH 4.6	37. 5.6% w/v Polyethylene Glycol 4000	37. 30% v/v
38. None	38. 0.09 M HEPES - Na pH 7.5	38. 1.26 M tri-Sodium Citrate dihydrate	38. 10% v/v
39. None	39. 0.085 M HEPES - Na pH 7.5	39. 1.7% v/v PEG 400, 1.7 M Sulfate	39. 15% v/v
40. None	40. 0.095 M tri-Sodium Citrate dihydrate pH 5.6	40. 19% v/v iso-Propanol, 19% w/v PEG 4000	40. 5% v/v
41. None	41. 0.085 M HEPES - Na pH 7.5	41. 8.5% v/v iso-Propanol, 17% w/v PEG 4000	41. 15% v/v
42. 0.04 M K-dihydrogen Phosphate	42. None	42. 16% w/v Polyethylene Glycol 8000	42. 20% v/v
43. None	43. None	43. 24% w/v Polyethylene Glycol 1500	43. 20% v/v
44. None	44. None	44. 0.1 M Magnesium Formate	44. 50% v/v
45. 0.16 M Zinc Acetate dihydrate	45. 0.08 M Sodium Cacodylate pH 6.5	45. 14.4% w/v Polyethylene Glycol 8000	45. 20% v/v
46. 0.16 M Calcium Acetate hydrate	46. 0.08 M Sodium Cacodylate pH 6.5	46. 14.4% w/v Polyethylene Glycol 8000	46. 20% v/v
47. None	47. 0.08 M Sodium Acetate trihydrate pH 4.6	47. 1.6 M Ammonium Sulfate	47. 20% v/v
48. None	48. 0.08 M Tris Hydrochloride pH 8.5	48. 1.6 M Ammonium dihydrogen Phosphate	48. 20% v/v
49. 0.8 M Lithium Sulfate monohydrate	49. None	49. 1.6% w/v Polyethylene Glycol 8000	49. 20% v/v
50. 0.4 M Lithium Sulfate monohydrate	50. None	50. 12% w/v Polyethylene Glycol 8000	50. 20% v/v

### Detergent Screen I

1. 0.08 mM C12E9
2. 0.11 mM C12E8
3. 0.17 mM n-Dodecyl-b-D-maltoside
4. 0.20 mM Sucrose monolaurate
5. 0.56 mM CYMAL-6
6. 0.90 mM TRITON X-100
8. 1.40 mM Deoxy BigChap
7. 1.00 mM CTAB
9. 1.80 mM n-Decyl-b-D-maltoside
10. 2.00 mM LDAO
11. 2.40 mM CYMAL-5
12. 4.00 mM ZWITTERGENT 3-12
13. 6.50 mM Nonyl-b-D-glucoside
14. 9.00 mM 1-S-octyl-b-D-thioglucoside
15. 10.4 mM DDAO
16. 19.5 mM HECAMEG
17. 24.4 mM n-Octanoylsucrose
18. 30.0 mM Heptyl-b-D-thioglucoside
19. 24.5 mM n-Octyl-b-D-glucoside
20. 34.5 mM CYMAL-3
21. 35.0 mM C-HEGA-10
22. 40.0 mM ZWITTERGENT 3-10
23. 79.0 mM MEGA-8
24. 250.0 mM n-Hexyl-b-D-glucoside

## Abbreviations

amu	atomic mass units
APS	ammonium peroxodisulfate
ATP	adenosine triphosphate
bp	base pairs
calc	calculated
CBD	chromophore binding domain
cry	cryptochrome
DBD	DNA binding domain
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
EDTA	ethylenediaminetetraacetate
fig.	figure
FR	far-red light
HKLD	histidine kinase-like domain
HPT	histidine phosphotransferase
Hrs	hours
HSQC	heteronuclear single quantum coherence
IEF	isoelectric focussing
IMAC	immobilized metal affinity chromatography
IPTG	isopropyl $\beta$ -D-thiogalactopyranoside
kD	kilodalton
LB	Luria Bertani broth
LDAO	N,N-dimethyldodecylamin-N-oxid
MAD	multiple anomalous wavelength dispersion
MIR	multiple isomorphous replacement
MR	molecular replacement
NMR	nuclear magnetic resonance
obs	observed
OD	optical density
ORF	open reading frame
PAS	Per-Arndt-Sim
PCB	phycocyanobilin
PCR	polymerase chain reaction
PEG	polyethylene glycol
Pfr	phytochrome, far-red light absorbing form
phot	phototropin
phy, PHY	phytochrome
pl	isoelectric point
Pr	phytochrome, red light absorbing form
PYP	photoactive yellow protein
pyp	photoactive yellow protein
P $\Phi$ B	phytochromobilin
R	red light

Rcp(s)	response regulator of cyanobacterial phytochrome (plural)
REC	receiver domain
RNA	ribonucleic acid
rpm	rounds per minute
RR(s)	response regulator (plural)
RT	room temperature
SAR	specific absorption ratio
SDS	sodium dodecylsulfate
SeMet	L-selenomethionine
TB	terrific broth
Tris	tris(hydroxymethyl)aminoethane
TWEEN	polyoxyethylenesorbitan
UV	ultraviolet
$\lambda_{\max}$	wavelength of the absorption maximum

## Danksagung

Die vorliegende Arbeit wurde unter Anregung von Prof. Wolfgang Gärtner am Max Planck Institut für Biochemie, Martinsried, in der Arbeitsgruppe von Prof. Dieter Oesterhelt durchgeführt. Ein Großteil dieser Arbeit wurde durch großzügige Förderung des Sonderforschungsbereichs 533, "Lichtinduzierte Dynamik von Biopolymeren" ermöglicht. Messungen von Röntgenbeugungsdaten erfolgten u.a. am Deutschen Elektronensynchrotron DESY, Hamburg, in der Gruppe von Dr. Hans Bartunik. Ihnen allen möchte ich meinen Dank ausdrücken.

Im besonderen gilt mein aufrichtiger Dank:

Herrn Prof. Wolfgang Gärtner, für die interessante Themenstellung, die überaus konstante Bereitschaft zur Diskussion, alle Unterstützung und Förderung und die mir anvertraute Freiheit zur wissenschaftlichen Entfaltung

sowie

Herrn Prof. Dieter Oesterhelt, für die Möglichkeit, im anregenden, wissenschaftlichen Umfeld seiner Arbeitsgruppe tätig zu sein und für sein Interesse und die großzügige Unterstützung

Frau Rita Wiemayer, für ihre Hilfsbereitschaft und liebevolle Unterstützung, ihr sonniges Gemüt und für bereichernde Konversationen und viel Lachen in der Teeküche

Frau Eleonore Hack, für viel Hilfe, und für alles was sie selbstlos und mit Herz so unbemerkt organisierte.

Frau Dr. Karin Rodewald, für die liebe und unermüdliche Leitung der Sequenzierabteilung

Allen anderen Mitglieder der Abteilung Membranbiochemie, für die freundliche und angenehme Atmosphäre und für alle Unterstützung die mir zukam, außerdem

Benjamin Quest und Stefan Schorling für eine meist freudvolle Laborgemeinschaft

Herrn Dr. Clemens Schäufler und Herrn Dr. Boris Bieger, für die hilfreiche und selbstlose Unterstützung bei röntgenkristallographischen Fragestellungen

Den fleißigen Studenten, die Ihren wertvollen Teil beigetragen haben, Richard Bögle, Thomas Hannich, Tobias Heckel, Elke Kolkmann, Frank Mathes, Andreas Traweger und besonders Julia Keller (kleine Häsin).

Walter Gmelin danke ich für die Vertrautheit, Freundschaft und lange Leidensgenossenschaft entlang eines "steinigen" Weges und für alle Hilfe

Andreas Schaer danke ich für die unendlich vielen Anregungen und den Besuch im Garten der Prinzessin Salome... ;-)

Meinen Freunden, die mir durchweg liebevoll und treu zur Seite standen und mich immer unterstützt haben, danke ich auf ewig. Sie sind die besten Freunde! Joe Sandmaier, Olli Scherer, Petra Müller, Petra Wollmann, Johanna Völkl (Jojo), Verena Zapletal (Vena), Claudia Macho, Mirjam Fruschella, Eva von Schaper und natürlich ... Mary J. Blige (I couldn't have done it without her!).

Schließlich danke ich auch meiner Schwester Amelie, meinem Bruder Sascha und Brigitte Anselment die mir halfen und mich unterstützten.

Am meisten danke ich meiner Mutter Vera, für ihre Liebe und dafür, dass sie mir all das ermöglicht hat, außerdem meinem Vater und Reza Memari. Sie standen immer an meiner Seite und ich Liebe sie über alles. Ihnen möchte ich meine Arbeit widmen.



## *Curriculum Vitae*

<b>Name</b>	Christian Gerhard Benda
Geburtsdatum	5. August 1969
Geburtsort	München
<b>Schulbildung</b>	
1976-80	Grundschule an der Nadistrasse, München
1980-1989	Lion-Feuchtwanger-Gymnasium, München
1989	Allgemeine Hochschulreife (Abitur)
<b>Auslandsaufenthalt</b>	
Okt. 1989 - April 1990	New York City, NY, USA, Sprachaufenthalt
<b>Ausbildung, Berufliche Erfahrung</b>	
Mai 1990 - März 1991	Romanische Philologie, Sprachkurse in Französisch und Spanisch
Sep. 1991 - Jan. 1992	Produktionsassistent bei Vision Film München, Filmproduktion
Nov. 1992 - Jan. 1999	Studium der Chemie an der Ludwig-Maximilians-Universität München, Studienschwerpunkt Biochemie
Mai 1998	Diplomhauptprüfung in Chemie
Juli 1998 - Jan, 1999	Diplomarbeit "Design und Herstellung von mRNA-Peptid-Fusionsmolekülen zur in vitro Selektion von cyclischen Peptiden" in der Arbeitsgruppe von Prof. Dr. Michael Famulok
<b>Promotion</b>	
März 1999 - März 2002	Doktorarbeit am Max-Planck-Institut für Biochemie bei Prof. Dieter Oesterhelt und Prof. Wolfgang Gärtner, Thema "Studies on Structure and Dynamics of Components of Phytochrom-mediated Light Signaling in Plants and Cyanobacteria"
seit April 2002	Proteros Biostructures GmbH, Martinsried