

Research article

Open Access

## Complementation of a phycocyanin-bilin lyase from *Synechocystis* sp. PCC 6803 with a nucleomorph-encoded open reading frame from the cryptophyte *Guillardia theta*

Kathrin Bolte<sup>†1</sup>, Oliver Kawach<sup>†1</sup>, Julia Precht<sup>†1</sup>, Nicole Gruenheit<sup>2</sup>, Julius Nyalwidhe<sup>3</sup> and Uwe-G Maier<sup>\*1</sup>

Address: <sup>1</sup>Philipps-Universität Marburg, Laboratorium für Zellbiologie, Karl-von-Frisch Str., D-35032 Marburg, Germany, <sup>2</sup>Heinrich-Heine Universität Düsseldorf, Institut für Botanik III, Universitätsstr. 1, D-40225 Düsseldorf, Germany and <sup>3</sup>Philipps-Universität Marburg, Laboratorium für Parasitologie, Karl-von-Frisch Str., D-35032 Marburg, Germany

Email: Kathrin Bolte - [kathrin.bolte@staff.uni-marburg.de](mailto:kathrin.bolte@staff.uni-marburg.de); Oliver Kawach - [oliverkawach@web.de](mailto:oliverkawach@web.de); Julia Precht - [julia.precht@web.de](mailto:julia.precht@web.de); Nicole Gruenheit - [nicole.gruenheit@gmx.de](mailto:nicole.gruenheit@gmx.de); Julius Nyalwidhe - [nyalwidh@staff.uni-marburg.de](mailto:nyalwidh@staff.uni-marburg.de); Uwe-G Maier\* - [maier@staff.uni-marburg.de](mailto:maier@staff.uni-marburg.de)

\* Corresponding author †Equal contributors

Published: 16 May 2008

Received: 14 December 2007

BMC Plant Biology 2008, 8:56 doi:10.1186/1471-2229-8-56

Accepted: 16 May 2008

This article is available from: <http://www.biomedcentral.com/1471-2229/8/56>

© 2008 Bolte et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** Cryptophytes are highly compartmentalized organisms, expressing a secondary minimized eukaryotic genome in the nucleomorph and its surrounding remnant cytoplasm, in addition to the cell nucleus, the mitochondrion and the plastid. Because the members of the nucleomorph-encoded proteome may contribute to essential cellular pathways, elucidating nucleomorph-encoded functions is of utmost interest. Unfortunately, cryptophytes are inaccessible for genetic transformations thus far. Therefore the functions of nucleomorph-encoded proteins must be elucidated indirectly by application of methods in genetically accessible organisms.

**Results:** Orf222, one of the uncharacterized nucleomorph-specific open reading frames of the cryptophyte *Guillardia theta*, shows homology to *slr1649* of *Synechocystis* sp. PCC 6803. Recently a further homolog from *Synechococcus* sp. PCC 7002 was characterized to encode a phycocyanin-β155-bilin lyase. Here we show by insertion mutagenesis that the *Synechocystis* sp. PCC 6803 *slr1649*-encoded protein also acts as a bilin lyase, and additionally contributes to linker attachment and/or stability of phycobilisomes. Finally, our results indicate that the phycocyanin-β155-bilin lyase of *Synechocystis* sp. PCC 6803 can be complemented *in vivo* by the nucleomorph-encoded open reading frame *orf222*.

**Conclusion:** Our data show that the loss of phycocyanin-lyase function causes pleiotropic effects in *Synechocystis* sp. PCC 6803 and indicate that after separating from a common ancestor protein, the phycoerythrin lyase from *Guillardia theta* has retained its capacity to couple a bilin group to other phycobiliproteins. This is a further, unexpected example of the universality of phycobiliprotein lyases.

### Background

Phycobiliproteins are subunits of the major accessory

light-harvesting complexes (LHC) of most cyanobacteria and red alga and are present in the thylakoid lumen of

cryptophytes as well. Covalently linked to the proteins are chromophore groups, the phycobilins [1,2]. These open tetrapyrrole rings are coupled to conserved cysteine residues via a thioether bond and are necessary for light harvesting and efficient energy flow [3]. Various phycobiliproteins, namely allophycocyanin, phycocyanin, phycoerythrin, phycoerythrocyanin, carry different numbers of bilin groups.

Attachment of bilins to phycobiliproteins is an enzymatically catalyzed reaction, which also occurs spontaneously, but at low efficiency [4]. Several bilin-attaching lyases are described. One of the dimeric enzymes encoded by *cpcE* and *cpcF* genes links the chromophore to the phycocyanin  $\alpha$ -subunit [4,5]. *PecE* and *pecF* genes encode the second known lyase, specific for the phycoerythrocyanin  $\alpha$ -subunit [6-8]. Recently Zhao and co-workers discovered that a CpeS-like protein functions as a phycocyanobilin-cysteine-beta84 lyase in *Anabaena* sp. PCC 7120, which was the first lyase identified for a  $\beta$ -subunit of a phycobiliprotein [9]. Another lyase specific for a  $\beta$ -subunit of a phycobiliprotein was found by Shen et al. [10]. They identified the gene product of *cpcT* to be a Cys- $\beta$ 153-phycocyanobilin lyase in *Synechococcus* sp. PCC 7002. Moreover, Zhao et al. reported the *Anabaena* sp. PCC 7120 CpeS1 as a "near-universal" lyase for cysteine-84-binding sites in cyanobacterial phycobiliproteins [11,12].

In most cyanobacteria and red algae phycobiliproteins are organized in multimeric complexes, called phycobilisomes [13-15]. Their antenna structure, located on the cytoplasmic surface of the thylakoid membrane, consists of various linker polypeptides and phycobiliproteins. Each phycobilisome is on its part a multimeric complex, composed of a core and several rod structures. Phycobilisomes can be subdivided according to their structure. The most common type in cyanobacteria, the hemidiscoidal one, consists of a tricylindrical core and six rods.

Allophycocyanin (AP,  $\lambda_{\max}$  = 650 nm) forms the core structure, connecting the phycobilisomes to the thylakoid membrane via linker proteins. Rods can be composed of three different phycobiliproteins: phycocyanin (PC,  $\lambda_{\max}$  = 617 nm) is located proximal to the core, whereas phycoerythrin (PE,  $\lambda_{\max}$  = 560 nm) and phycoerythrocyanin (PEC,  $\lambda_{\max}$  = 575 nm) are located distal to the core [16,17]. The phycobilisome rods of each organism differ in their phycobiliprotein composition. *Synechocystis* sp. PCC 6803 harbors hemidiscoidal phycobilisomes. PC, the only biliprotein in the rod structures in this organism, is composed of  $\alpha$ - and  $\beta$ -subunits. These subunits dimerize to heterodimers, assemble to hexameric ( $\alpha\beta$ )<sub>6</sub> discs, and are subsequently coupled to each other, as well as to the AP-core via linker proteins [18,19]. Depending on their location (in core or rods), and their molecular mass,

linker proteins are divided into four groups [20,21]. Beside their main function of mediating the assembly and stability of the phycobilisomes, linker proteins also promote energy transfer towards the reaction centres [20].

*Guillardia theta* is a cryptophyte possessing phycoerythrin as a phycobiliprotein. The  $\beta$ -subunit is encoded on the plastid genome [22], whereas the phycoerythrin  $\alpha$ -subunits are encoded by a nuclear-located gene family. In the latter case, the genes encode preproteins containing a tripartite topogenic signal responsible for the translocation across five biological membranes [23]. Because a wide range of genomic data exists from this unicellular phototrophic organism, existing knowledge can be used to reconstruct the biochemistry of these organisms. The elucidation of protein functions encoded by open reading frames in the nucleomorph genome of *Guillardia theta* is of special interest, as this genome is minimized and should therefore encode only essential proteins. After analyzing the nucleomorph genome data, Orf222 was identified as being homologous to a number of proteins including Slr1649 from *Synechocystis* sp. PCC 6803 and CpcT from *Synechococcus* sp. PCC 7002. Because cryptophytes are inaccessible to genetic manipulations, we created a *slr1649*-loss-of-function strain of *Synechocystis* sp. PCC 6803 and complemented this strain with the nucleomorph-encoded orf.

The generated Slr1649 loss-of-function mutant generally has characteristics conductive with the description by Shen et al. for a *cpcT* knock-out in *Synechococcus* sp. PCC 7002 [10]. Nevertheless, additional effects in the *slr1649* knock-out mutant of *Synechocystis* sp. PCC 6803 were identified in respect to linker proteins within the phycobilisomes of the mutant. Complementation of *slr1649* with the nucleomorph-specific *orf222* indicated that the cryptophytic protein, although having originated from an organism using phycoerythrin as accessory pigment, attaches a bilin to the position Cys- $\beta$ 155 of phycocyanin in the cyanobacterium.

## Results

### In silico analyses

After analyzing the nucleomorph genome of the cryptophyte *Guillardia theta*, we identified an open reading frame (*orf222*) with a high degree of similarity to cyanobacterial genes (Table 1), that encode soluble proteins possessing a DUF1001 domain. Alignments of the cryptophyte sequence with these cyanobacterial sequences indicated that *orf222* should encode an additional transit peptide as shown by a N-terminal extension. Further orfs with homology to *orf222* and the cyanobacterial homologs are additionally present in the nuclear genome of red alga [24]. In higher plants, i.e. *Arabidopsis thaliana*, orfs with some similarity are also present [25]. Even the

**Table 1: Tabular comparison of prokaryotic and eukaryotic homologous of Orf222**

Organism		PBP	Homologous of Orf222	Length (aa)	CpeT homolog	Slr1649 homolog	Genomic context of the encoded genes
<i>Synechocystis</i>	S. sp. PCC 6803	PC	Slr1649	196			<i>slr1648/ssr2754/slr1649</i>
<i>Crocospaera watsonii</i>	WH 8501	PE	CwatDRAFT_4238	196		x	<b><i>cwatDRAFT_4238</i></b> / <i>cwatDRAFT_4297</i>
		PC	CwatDRAFT_0664	215	x		<i>cpeA/cpeT/cpeY</i>
			CwatDRAFT_5720	149	x		<i>cpeS/cpeT/cpeR</i>
<i>Nostoc</i>	<i>Nostoc punctiformes</i> PCC 73102	PEC	Npun02004130	197		x	<i>thrC/npun02004130/npun02004132</i>
		PC	Npun02004123	209	x		<i>cpeS/cpeT/cpeR</i>
	<i>Anabaena</i> sp. PCC7120		Npun02007740	201	x		-
			All5339	199		x	-
			Alr0647	198	x		-
<i>Anabaena variabilis</i> ATCC 29413		Ava_2579	199		x	-	
		Ava_4579	198	x		-	
<i>Thermosynechococcus elongatus</i>	BP-I	PC	Tlr2156	196		x	<i>tlr2154/hemD/tlr2156</i>
<i>Synechococcus</i>	<i>S. elongates</i> PCC 6301	PC	Syc0738_d	197		x	<b><i>syc0738_d</i></b> / <i>syc0739_d/ruvC</i>
		PE	Syc0764_d	197	x		-
	<i>S. sp.</i> PCC 7002		CpcT	199			from [10]
		<i>S. sp.</i> CC 9311		Sync_0487	196		x
	<i>S. elongates</i> PCC 7942		Sync_0509	208	x		<i>cpeC/sync_0512/cpeD-1/cpeS/sync_0509</i>
			Synpcc7942_0772	197	x		-
	<i>S. sp.</i> CC 9605		Synpcc7942_0800	197		x	<b><i>synpcc7942_0800</i></b> / <i>synpcc7942_0799/ruvC</i>
			Syncc9605_0440	204	x		<i>cpeS/cpeT/cpeR</i>
	<i>S. sp.</i> WH 8102		Syncc9605_0419	208		x	<i>cpcB/cpcA/syncc9605_0419/phycoerythrin-lyase</i>
			SYNW2024	197		x	<i>rpcB/rpcA/synw2024/phycoerythrin-lyase</i>
<i>S. sp.</i> CC9902		SYNW2003	204	x		<i>cpeC/mpcD/cpeE/cpeS/cpeT/cpeR</i>	
		Syncc9902_1910	200		x	<i>cpcB/cpcA/9902_1910/phycoerythrin-lyase</i>	
<i>Trichodesmium erythraeum</i>	IMS101	PC	Tery_0543	195		x	<i>cyp/cyp/tery0543/transposase</i>
		PE	Tery_0979	209	x		<i>cpeZ/cpeT/cpeF</i>
<i>Calothrix</i>	PCC 7601	PC PE	CpeT	207	x		<i>cpeS/cpeT/cpeR</i>
<i>Gloeobacter violaceus</i>	PCC 7421	PE	Glrl182	202			<i>apcD/glrl182/cpcB/cpcA</i>
		PC	Glrl193	203	x		<i>cpeS/cpeT</i>
			Glrl538	183	x		-
<i>Prochlorococcus marinus</i>	MIT 9211	PE	P9211_07167	192	x		<i>cpeS/cpeT</i> All adjacent to the $\alpha$ - and $\beta$ - subunit of phycoerythrin
	CCMP1375		Orf195	195	x		<i>cpeS/cpeT</i>
	MIT 9313		PMT1678	239	x		<i>cpeS/cpeT</i>
	NATL2A		PMN2A_1676	199	x		<i>cpeS/cpeT</i>
	SSI20		Pro0342	195	x		<i>pro0347/pucC/ppeC/pro0344/cpeS/cpeT</i>

**Table 1: Tabular comparison of prokaryotic and eukaryotic homologous of Orf222 (Continued)**

<i>Guillardia theta</i>	PE	Orf222	222
<i>Cyanidioschyzon merolae</i>	PC	CMK263C	263
Bacteriophage S-PM2	host = Synechococcus	S-PM2p215	175
<i>Oryza sativa</i>		LOC_Os11g3216 0	275
<i>Arabidopsis thaliana</i>		AT5G51020	269

Cyanobacterial and eukaryotic homologous of Orf222 were obtained by blast search analysis of the NCBI database, as well as the JGI and Cyanobase. The (-) indicates that no adjacent genes in the same orientation were detectable. Orf222 homologous genes are in bold. PBP = phycobiliproteins (except of allophycocyanin), PC = phycocyanin, PE = phycoerythrin, PEC = phycoerythrocyanin.

bacteriophage S-PM2, which infects *Synechococcus* strains, encodes a homolog of *orf222* [26] (Table 1).

The number of Orf222 homologues in cyanobacteria varies in several species and does not correlate with the number of phycobiliproteins. Nevertheless, there is a strong tendency to express more than one species of Orf222 homolog in organisms containing multiple types of phycobiliproteins in the rods (Table 1). Based on amino acid sequence alignments and phylogenetic networks, four monophyletic groups can therefore be assigned (Fig. 1). Two of them resemble CpeT-like proteins (phycoerythrin operon protein); the other two groups harbor members of Slr1649-like type. Neither the *Guillardia theta* sequence nor any other eukaryotic sequences can be assigned to any one of the four monophyletic groups.

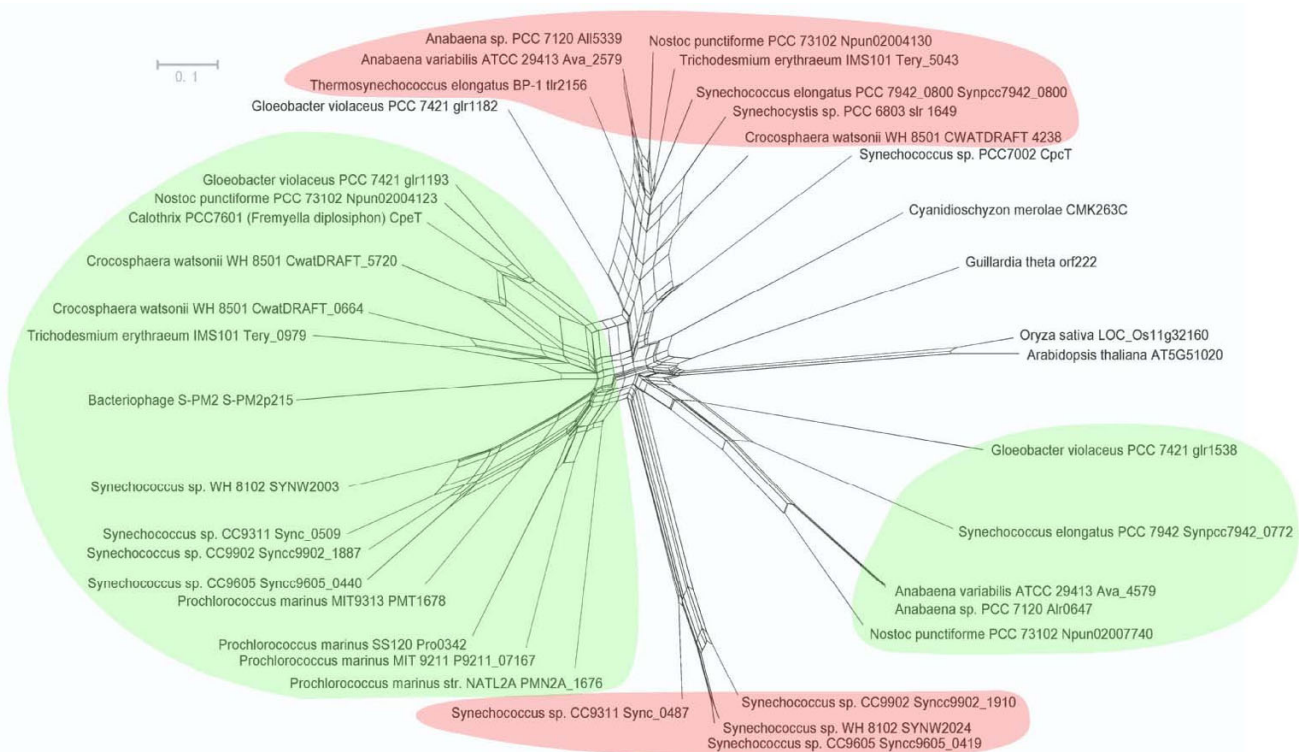
Additionally, clear affiliations of the *Gloeobacter violaceus* PCC 7421 (glr1182) and *Synechococcus* sp. PCC 7002 (CpcT) sequences can not be extrapolated. With the exception of the *Prochlorococcus* species, at least one member of the Slr1649-like group is present in all cyanobacteria investigated to date. CpeT-like proteins were only detected in cyanobacteria encoding phycoerythrin and/or phycoerythrocyanin. Although the proteins of both groups seem to have the same function, further investigations on the corresponding genes relevant in the genomic context revealed a noticeable difference. Unlike the genes of the *cpeT*-group, the *slr1649*-group is by far less conserved in its genomic localization (Table 1). Except for *Nostoc* sp. PCC 7120, *Anabaena variabilis* ATCC 29413 and *Trichodesmium erythraeum* IMS 101, the localization of the homolog gene is always downstream of *cpeS*. In few cases it is followed by *cpeR*. On the other hand, genes for the *slr1649*-group are rather randomly distributed in the investigated cyanobacterial genomes (Table 1).

### Generation of a *slr1649* knock-out strain

We used *Synechocystis* sp. PCC 6803 as a model organism and created first a *slr1649* knockout strain ( $\Delta$ *slr1649*) by inserting a kanamycin resistance cassette into the *slr1649* open reading frame via homologous recombination. The generated homozygous knock-out mutant shows identical features described in Shen et al. [10]. Just like the characterized *cpcT* mutant in *Synechococcus* sp. PCC 7002, our knock-out mutant contains a decreased level of phycocyanin up to 60% and a resulting pale green phenotype. The knock-out cells produce smaller phycobilisomes, which could be the cause of their different migration behaviour in sucrose density gradients in comparison to wild type phycobilisomes. Furthermore, isolated phycobilisomes showed a red-shifted absorbance maxima and a slightly smaller apparent molecular mass in the  $\beta$ -subunit of phycocyanin on SDS-PAGE (data not shown). After the digestion of purified phycocyanin with formic acid and a phycocyanobilin addition assay, Shen et al. concluded after digestion that the *cpcT* gene from *Synechococcus* sp. PCC 7002 encodes a bilin lyase responsible for the attachment of phycocyanobilin to Cys-153 on the  $\beta$ -subunit of phycocyanin [10]. The same is most likely true for the *Synechocystis* homolog Slr1649 due to the high homology and the similar phenotype between the two knock-out mutants. Thus, Slr1649 is thought to attach a bilin group to the homolog position Cys-155 of  $\beta$ -phycocyanin in *Synechocystis* sp. PCC 6803.

### Isolation and analysis of phycobilisomes

By isolating and analysing phycobilisomes from our knock-out mutant, we observed one additional feature besides the one already known from the characterization of the *Synechococcus* sp. PCC 7002 homolog. As shown in Fig. 2A two linker polypeptides, CpcC2 and CpcD, encoded by the genes *cpcC2* and *cpcD* respectively, are missing in the isolated phycobilisomes of the mutant but could be identified in the wild type during mass spectrometry analysis. The other linker polypeptides CpcC1



**Figure 1**

**NeighborNet (NNet) splits graph for 41 taxa.** Proteins sequences were aligned with MUSCLE. The initial alignment contained 307 sites including 191 gapped sites that were excluded from the analysis, leaving 116 amino acid sites for log determinant (LogDet) distance estimates with removal of invariant sites using the program LDDist. From this a Neighbor – net splits graph was constructed, which is visualized with Splitstree4. Highlighted are four monophyletic groups: Two of them resemble CpeT-like (phycoerythrin operon protein) proteins (highlighted in green) and two groups harbor members of the Slr1649-like type (highlighted in red). Not shown: The sequences of *Synechococcus elongatus* PCC 7942 Synpcc7942\_0800 and *Synechococcus elongatus* PCC 6301 Syc0738\_d are identical as well as the sequences of *Synechococcus elongatus* PCC 7942 Synpcc7942\_0772 and *Synechococcus elongatus* PCC 6301 Syc0764\_d and the sequences of *Prochlorococcus marinus* SS120 Pro0342 and *Prochlorococcus marinus* CCMP1375 orf195.

(*cpcC1*), CpcG1 (*cpcG1*), ApcC (*apcC*) and ApcE (*apcE*) are present in both strains.

Upon further investigations, the absence of the linker proteins was proven not to be a transcriptional effect. In exemplarily reverse transcription experiments for *cpcC2*, the presence of identical *cpcC2* transcripts was confirmed in both the mutant and the wild type strain by sequencing the obtained RT-PCR products (data not shown).

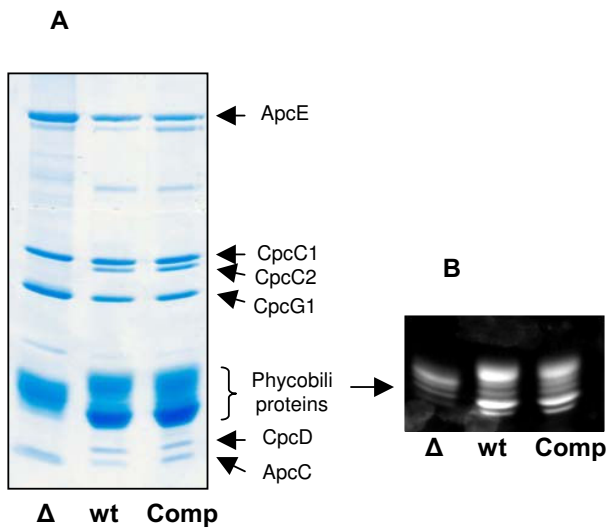
#### Complementation

In order to investigate if the nucleomorph-specific reading frame *orf222* from the cryptophyte *Guillardia theta* is able complement the effects of the *slr1649* loss-of-function, we integrated this potential gene without its putative transit peptide into the cyanobacterial genome of *Synechocystis* sp. PCC 6803. This simultaneously affected the reading frame of *slr1649* and its *cis*-acting upstream signals (Fig.

3A). In the complemented strain, *slr1649* is separated from its natural upstream region, generating a promoterless truncated gene, in which the translational initiator codon and the next two codons are no longer present in the reading frame. The loss of the *slr1649* gene product and the complete segregation of the mutation were shown by immunoblot experiments using polyclonal antibodies generated against Slr1649 (Fig. 3B). Here, cross-reactions of the antibody were shown to be present in the wild type but not in the mutant strain extract. Additional analysis of the complemented strain by RT-PCR and sequence analysis showed that the integrated cryptophytic *orf222* is transcribed (data not shown).

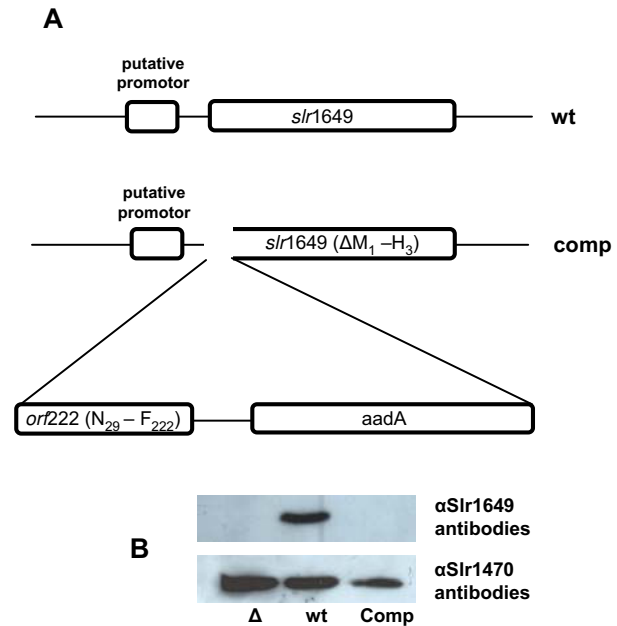
#### Characterization of the complemented strain

Interestingly the phenotype of the complemented strain is similar to that of the wild type strain, as indicated by the greenish color of the culture (data not shown). During



**Figure 2**  
**SDS-PAGE of phycobilisomes isolated from sucrose gradients.** (A) Coomassie staining of separated phycobilisome components from  $\Delta slr1649$  ( $\Delta$ ), wild type and complemented (Comp) strains. Both linker proteins (CpcC2 and CpcD), which are absent in  $\Delta slr1649$  cells, are present in the complemented strain, indicating a wild type phycobilisome structure. The linker proteins CpcC2 and CpcD are absent in the mutant strain ( $\Delta$ ). (B)  $Zn^{2+}$  staining of phycobilisomes separated in SDS-PAGE. The signal intensity of phycobiliproteins isolated from wild type and complemented (Comp) strains seems to be equal in contrast to the reduced signal of the  $\Delta slr1649$  mutant.

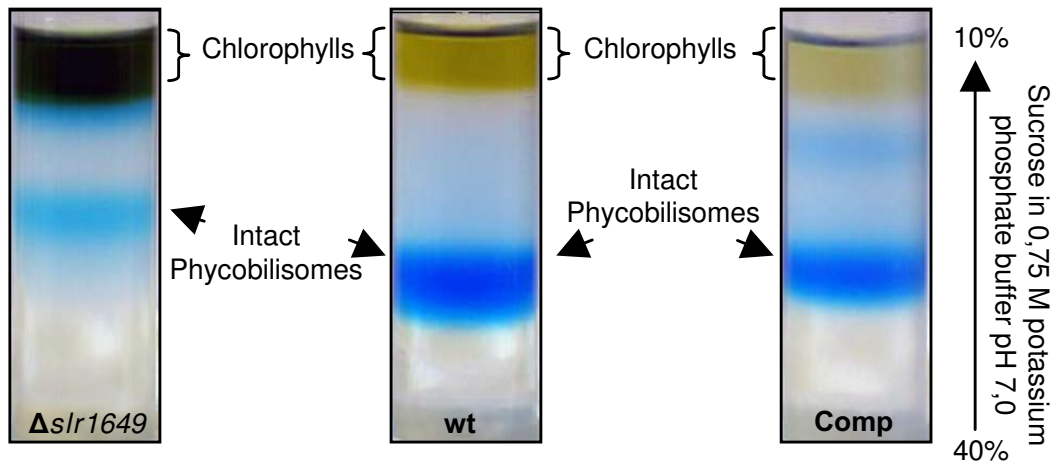
sucrose density-gradient separation of isolated phycobilisomes from the complemented strain, we noticed no difference in the migration behaviour of the prominent band in respect to the wild type strain in contrast to the migration behaviour of the knock-out mutant (Fig. 4). This indicates that the size of the phycobilisomes is identical in both strains and could indeed be confirmed by a proteome and mass spectrometry analysis. Here we showed that the missing linker proteins of the *slr1649* knock-out strain were present in the complemented strain (Fig. 2A). Additionally, there was no molecular mass shift in the  $\beta$ -subunit of phycocyanin of the complemented strain on SDS-PAGE visible. Further analyses revealed that the chromophore group, missing at positions Cys- $\beta$ 153 and Cys- $\beta$ 155 in the knock-out mutants of *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 respectively, most probably reappeared in the complemented strain, because  $Zn^{2+}$  stainings of phycobilisomes separated by SDS-PAGE showed an equal signal intensity of the phycobiliproteins of the wild type and complemented strain (Fig. 2B).



**Figure 3**  
**Complementation construction and control experiments.** (A) Schematic picture of the complementation construct. Schematic depiction of the construct used for the complementation of *slr1649* with the cryptophytic *orf222*. The upper figure displays the wild type situation. The lower figure shows the insertion site of *orf222*, without its putative transit peptide and the *aadA* gene into *slr1649* and its *cis*-acting upstream signals. (B) Immunoblot with Slr1649 (upper) and Slr1470 (lower) specific antibodies. Cells from  $\Delta slr1649$ , wild type and complemented strains were disrupted and the protein extracts were separated by SDS-PAGE. Neither in the fraction of  $\Delta slr1649$  cells nor in the fraction of the complemented strain were signals of the Slr1649 antibody detectable. Specific polyclonal Slr1470 antibodies were used as a loading control and clear signals at the expected size were obtained in all three protein fractions.

To clarify this we digested isolated phycobiliproteins with formic acid. In doing so, CpcB is cleaved at a single site while all other phycobiliproteins remain unaffected. The expected sizes for fluorescent fragments are 15.36 kDa with a chromophore group at position Cys- $\beta$ 84 and 2.78 kDa for the fragment with the chromophore group at position Cys- $\beta$ 155. As shown in Figure 5, these expected fragments were obtained. In addition to a signal at 15.36 kDa, a signal at 2.78 kDa was detected in the lane containing wild type strain protein and protein from the complemented strain but not in the one containing  $\Delta slr1649$ . These data confirmed that the chromophore group, missing in the knock-out mutant, is present in the complementation. This together with the identification of





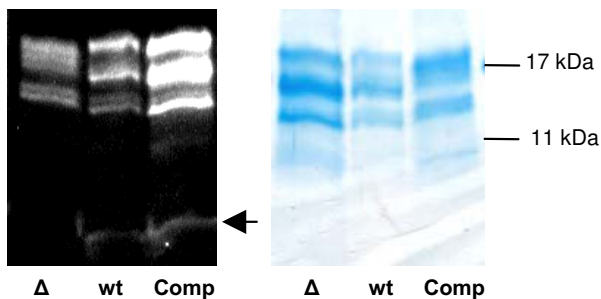
**Figure 4**  
**Isolation of intact phycobilisomes.** Phycobilisomes were isolated as described in Material & Methods. After 16 h centrifugation the phycobilisomes became visible as clear blue bands in the gradient. The upper layer contained chlorophylls. The phycobilisomes of  $\Delta slr1649$  cells had a diminished migration compared to the wild type ones, whereas the phycobilisomes from the complemented strain (Comp) had a migration equivalent to the wild type.

the linker protein spectrum in the complemented strain indicated a wild type phycobilisome structure.

**Discussion**

Cryptophytes are important organisms for several reasons. In terms of cell biology, their complex compartmentalization is of major interest, because several plasmas and genomes coexist in these organisms, which can be traced back to either a prokaryote or a eukaryote [27]. One of the hallmarks of cryptophytes is the remnant of a second nucleus, which originated by the reduction of the cell nucleus of an engulfed phototrophic eukaryote by

another eukaryotic cell [28]. This compartment, the nucleomorph, is minimized in its coding capacity and expresses – in the case of *Guillardia theta* – only approximately a tenth of that of the *E. coli* K12 genome [29]. The reduced coding capacity leads to the impression that the genes are still present in the nucleomorph may encode important functions. Thus, we are interested in addressing the functions of proteins encoded by the nucleomorph. However, due to the lack of a method of transfecting cryptophytes, we are studying homologs of the nucleomorph genes and their encoded proteins in genetically accessible organisms in order to identify the functions of the cryptophytic proteins indirectly. One of the best-studied and genetically accessible cyanobacterium is *Synechocystis* sp. PCC 6803.



**Figure 5**  
**Proteolytic digestion of phycobilisomes.** (A) Digestion of phycobilisomes with formic acid. The arrow indicates the resulting fragment after  $Zn^{2+}$  stain at 2.78 kDa. There are also several signals at 17–20 kDa which refer to the unaffected  $\alpha$ -subunit of phycocyanin and the  $\alpha$ - and  $\beta$ -subunit of allophycocyanin.

Orf222 is one of the uncharacterized nucleomorph-specific open reading frames, for which homologs are present in many cyanobacteria. Analysis of the contribution of this gene within different organisms indicated that a clear correlation between *orf222*-homolog genes and phycobiliproteins is present, because at least one *orf222* homolog is encoded in all organisms expressing phycobiliproteins, including red alga. Phylogenetic studies demonstrated that homologs of the *orf222* gene can be classified into the following four groups (Fig. 1): Slr1649-like a, Slr1649-like b, CpeT-like a and CpeT-like b. Because the method for network construction as well as sampling in our studies is different from that of a recently presented phylogeny [10], it is not surprising that slightly different affiliations are resolved. However, our network

corrects erroneous affiliations and indicates uncertainties of the basal grouping. This may be seen in the position of the bacteriophage sequence, which is in the network presented here in the neighbourhood of the bacteria they infect and not in the same branch as the cryptophyte sequence.

Despite the high degree of homology, the members of Slr1649-like and CpeT-like groups differ in the genomic context of the corresponding genes (Table 1). Members of the CpeT group are predominantly localized in the phycoerythrin associated linker protein operon [30,31] next to the *cpeS* gene. In some cases, even the *cpeR* gene is localized directly downstream of *cpeT*. Because operon structures connect functionally related genes in many cases, CpeT could be an either structurally or functionally moiety of the phycobilisome and it has been shown to be responsible for the attachment of a bilin group to a specific site from  $\beta$ -phycocyanin [10]. It is remarkable that a congruent distribution of members of the Slr1649-groups is not visible, because the genes seem to be localized randomly throughout different genomes. Interestingly, *slr1649*-homologs exist in some higher plants such as *Oryza sativa* and *Arabidopsis thaliana* (Fig. 1). The encoded proteins of these land plants are characterized by a DUF1001 domain as well, but obviously have paralogous functions, since the *Arabidopsis thaliana* homolog seems to be required for plastid division [25]. It is also suggested to play an important role in cell differentiation and the regulation of the cell division plane in plants [25]. The same could be true for the copy of the bacteriophage S-PM2, but seems to be unlikely since this phage infects different *Synechococcus* strains and its resource of the homolog may be the result of a selective advantage.

The homozygous knock-out mutant  $\Delta slr1649$  in *Synechocystis* sp. PCC 6803 showed features identical to a *cpcT* knock-out mutant from *Synechococcus* sp. PCC 7002 described in Shen et al. [10]. Here, the same pale green phenotype and a reduced phycocyanin content, resulting from a missing bilin group in phycobilisomes, was created by knock-out of *cpcT*, homolog to *slr1649* homolog in this cyanobacterium. This indicates that the lyase function of the homologous proteins of *Synechococcus* sp. PCC 7002 and *Synechocystis* sp. PCC 6803 is comparable.

Nevertheless, we obtained one additional, not described feature in the *Synechocystis* sp. PCC 6803 knock-out mutant. Two linker proteins, CpcC2 and CpcD, were missing from the phycobilisomes in the knock-out mutant  $\Delta slr1649$ . CpcD is a small linker (10 kDa) located at the distal tip of rods, possibly functioning as a rod terminating factor [32]. The CpcC2 rod linker (30 kDa) connects the most distal located phycocyanin discs [33]. Both genes are located in the phycocyanin operon from which

they are co-transcribed with the phycocyanin subunits and the *cpcC1* linker gene [33]. A transcriptional effect causing the loss of the linker proteins appears to be very unlikely, because the  $\alpha$ -subunit, the  $\beta$ -subunit and CpcC1 linker are present, although the CpcD and the CpcC2 linker are simultaneously absent. This is indicative of our finding that the *cpcC2* gene is indeed transcribed in the mutant as indicated by reverse transcription experiments (data not shown). Therefore, the deficit of the two linker proteins in mutant phycobilisomes is a post-transcriptional effect. However, we can not rule out that a decreased stability of phycobilisomes caused by the altered  $\beta$ -phycocyanin may be the reason for the lack of the two linker proteins in our preparations. In any case, the lack of the linker proteins is a molecular marker for the loss of lyase function, which may be interested to be studied in *Synechococcus* sp. PCC 7002 [10] as well.

*Guillardia theta*, the cryptophyte on which we are primarily focusing expresses a homolog of *slr1649* in association with phycoerythrin. Phycobiliproteins are located in the thylakoid lumen and apparently not organized in phycobilisomes in cryptophytes. Because the cryptophyte *Guillardia theta* uses phycoerythrin and not phycocyanin as an accessory pigment for photosynthesis, one might not expect that the putative cryptophytic lyase is able to complement the one of *Synechocystis* sp. PCC 6803. Surprisingly, the complemented strain showed wild type phycobilisomes structures as shown by the correct attachment of chromophore groups and the linker protein spectrum. Thus, Orf222 from the cryptophyte is able to complement the loss-of-function of Slr1649, indicating that the cryptophytic phycoerythrin lyase has still retained the capacity to couple a bilin group to  $\beta$ -phycocyanin, even after the progenitor of both classes of proteins evolved into apparently paralogous ones. However, a pleiotropic function of a biliprotein lyase with a specificity for phycobilin:cysteine-84 was recently shown *in vitro* for CpeS1 from *Anabaena* PCC 7120 [11], implicating that a multiplicity of proteins like the cryptophytic phycobilin:cysteine- $\beta$ 155 lyase has the capacity to couple bilins to homologous positions in a variety of phycobiliproteins.

## Conclusion

Loss-of function of a bilin lyase leads to a variety of effects in phycobilisome structure. This is already shown for a *cpcT* mutant in *Synechococcus* sp. PCC 7002 and could be confirmed by the generation of a *slr1649* knock-out mutant in *Synechocystis* sp. PCC 6803, homolog *cpcT*. One additional feature, the lack of two distal linker proteins, fits with the already known altered phycobilisome structure and may be the reason for the decreased phycocyanin content in mutant missing the bilin lyase.



Loss of Slr1649 was complemented *in vivo* by the homolog Orf222, which is encoded by the tiny vestigial nucleus of the eukaryotic endosymbiont from the cryptophyte *Guillardia theta*. Thus, Orf222 is supposed to be a phycoerythrin-bilin lyase in cryptophytes. Despite having originated from an organism using phycoerythrin as its accessory pigment, the protein still has the capacity to couple a chromophore group to the  $\beta$ -subunit of phycocyanin, indicating the functional universality of bilin lyases on the one hand and demonstrating the importance of nucleomorph-encoded cellular functions on the other.

## Methods

### Cell Culture

*Synechocystis* sp. PCC 6803 strains, wild type,  $\Delta$ slr1649 and the complemented strain, were grown at 30°C in Erlenmeyer flasks containing BG-11 media [34] with gentle swirling under standard light conditions (70  $\mu$ E) and atmospheric CO<sub>2</sub> levels. For growth on plates, BG-11 medium was supplemented with 1% Agar. Plates were incubated under the same conditions as liquid cultures.

### Construction of the $\Delta$ slr1649 Mutant

Two pairs of primers were used to amplify the flanking regions of the knock-out construct: 1649a\_f (5'-GGT TAC TGC TCG AGG CGC ATC A-3') and 1649a\_r (5'-GGA CGG CAA GGG ATC CTA TCT GG-3') generate fragment slr1649a, 1649b\_f (5'-GGA CGG CAA GGG ATC CTA TCT GG-3') and 1649b\_r (5'-CAG AAA TTG CCG CGG CCA ATC TC-3') fragment 1649b. Both were ligated into the pGEM-T vector (Promega, Mannheim) and after verification of the sequence, transferred into the pBluescript II SK (Stratagene, Amsterdam) vector. *Escherichia coli* strain MRF<sup>+</sup> XL-1 blue was used as plasmid host for cloning steps.

Using the *Bam*HI restriction site (inserted by the primer 1649a-r and 1649b\_f), a kanamycin resistance gene was cloned between the two fragments resulting in plasmid p $\Delta$ 1649. After transforming into wild type *Synechocystis* sp. PCC 6803 cells with this plasmid, transformants were selected on BG-11 agar plates supplemented with kanamycin (5  $\mu$ g/ml starting concentration). Kanamycin resistant clones were transferred to BG-11 liquid media. A homozygous culture was achieved by increasing kanamycin concentrations (50  $\mu$ g/ml final concentration). Complete knock-out was confirmed via Southern blot analysis.

### Construction of the Complemented Strain

Nucleotide sequence of orf222 was amplified from *Guillardia theta* DNA without its putative transit peptide by using the primers 222komp2\_f (5'-CAT ATG AAT TAA AAC CAA TCC TTA ATT G -3') and 222komp\_r (5'-GTT AAA ATT AAA TGA ATT CTA ATA A-3'). Two pairs of primers were used to amplify the flanking regions: 1649a\_f (5'-

GGT TAC TGC TCG AGG CGC ATC A-3') and 1649komp1\_r (5'-CAA TAA CTA CAT ATG TCC CAT TCC-3') generated the fragment Compa, which includes the upstream region for slr1649, 1649komp2\_f (5'-TTT ATG TCG AAT TCC ACT GAT C-3') and 1649b\_r (5'-GAG ATT GGC CGC GGC AAT TTC TG-3') generated fragment Compb. All three fragments were ligated into the pGEM-T vector (Promega, Mannheim) and after verification of the sequence, transferred into the pBluescript II SK (Stratagene, Amsterdam) vector using different restriction sites inserted by the primers leading to a precursor construct. By using *Eco*RI restriction sites, a spectinomycin cassette was cloned between the two fragments Compa/orf222 and Compb resulting in plasmid pComp222  $\Delta$ slr1649. After transforming the *Synechocystis* sp. PCC 6803 wild type strain with this plasmid, transformants were selected on BG-11 agar plates supplemented with spectinomycin (5  $\mu$ g/ml starting concentration). Spectinomycin resistant clones were transferred to BG-11 liquid media. A homozygous culture was achieved by increasing spectinomycin concentrations (30  $\mu$ g/ml final concentration).

### Nucleic Acid Analysis

*Synechocystis* sp. PCC 6803 cells were collected by centrifugation of 5 ml cell culture at 3200  $\times$  g. For DNA isolation, the pellet was resuspended in 400  $\mu$ l TE buffer pH 7.0. After addition of breaking buffer (10% sodium dodecyl sulfat (w/v), 5% sodium lauryl sulfat (w/v)), 200  $\mu$ l glass beads (0.2 mm diameter) and 400  $\mu$ l phenol, cells were lysed by vortexing the suspension three times for 10 s. The suspension was then centrifuged at 12 000  $\times$  g and the resulting upper phase transferred to a new cup. This sample was treated twice with phenol-chloroform-isoamylalcohol (25:24:1) and centrifuged as before. By adding 1/10 Vol. NaAc pH 4.8 and two Vol. 96% ethanol, the DNA was precipitated for 1 h at -20°C. Afterwards, an additional washing step with 70% ethanol was performed. The pellet was dried and resuspended in H<sub>2</sub>O.

RNA was isolated from *Synechocystis* cells with Trizol<sup>®</sup> (Invitrogen, Karlsruhe) according to the manufacturer protocol. Northern Blot and Southern Blot analysis were performed according to standard protocols (Sambrook). Probes were constructed using the PCR DIG Probe Synthesis Kit (Roche, Mannheim).

### Antibody Generation and Purification

To generate an antibody against Slr1649 we used the primers ex1649\_f (5'-GGA TCC TTA TGT CCC ATT CCA CTG-3') and ex1649\_r (5'-CTC GAG GCT GGC TAA AAA CTA ACT-3') to amplify the slr1649 gene, which was finally cloned in the pGEX-5X-3 vector (GE Healthcare Biosciences). After overexpression and purification of the Slr1649 GST fusion protein, immunization steps were executed by the Eurogentec company (Seraing).

The IgG fraction was purified from serum by protein A sepharose beads (GE Healthcare Biosciences).

#### **Isolation of Phycobilisomes**

Phycobilisome isolation was performed according to Gray et al. [35]. Cells were collected by centrifugation at 5000 rpm for 10 min at room temperature. After an additional washing step with BG-11 media, cells were resuspended in 0.75 M potassium-phosphate buffer pH 7.0 (PPB), containing a protease inhibitors cocktail (PIC, 2 mg/ml Antipain, 5 mg/ml Chymostatin, 2 mg/ml Aprotinin, 5 mg/ml Trypsin-Inhibitor, 2 mg/ml Pepstatin, 5 mg/ml Leupeptin, 1 mg/ml Elastatinal and 2 mg/ml Na<sub>2</sub>EDTA in HEPES/KOH. Final concentration 200 µg/ml Inhibitor) and afterwards broken by two passes through a French press (Aminco) at 124 MPa. The lysates were incubated with Triton X-100 (2%) for 15 min at room temperature and subsequently centrifuged at 20 000 rpm for 1 h to pellet unbroken cells and membrane debris. The supernatant was immediately loaded on a 10%–40% linear sucrose gradient, solved in PPB and centrifuged at 18 000 rpm for 16 h at 15 °C.

#### **SDS-PAGE**

Standard SDS-PAGE was performed with an Hoefer SE 250 apparatus (83 mm × 101 mm, 0,75 mm thick) or a custom made system (250 mm × 150 mm and 1,0 mm thick) using the Laemmli buffer system [36]. The polyacrylamide content in the separating gel was a gradient of 10% to 15%. The stacking gel contained 6% polyacrylamide. To achieve a better resolution of polypeptides with masses less than 15 kDa, the SDS-Tricine gel system was used [37]. Staining of gels was generally carried out with Coomassie brilliant blue G-250 dissolved in solution A (2% phosphoric acid v/v, 10% (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> w/v) and methanol (40:9:1). To visualize the bilin carrying proteins gels were incubated in a 0.2 M ZnSO<sub>4</sub> solution [38,39] and highlighted with UV in a transilluminator (Bio-Rad).

#### **Formic Acid Cleavage**

Phycobilisomes were precipitated with Methanol/Chloroform [40] and resuspended in cleavage buffer. Cleavage was done according to Piszkiwicz et al. [41]. 30 µg of phycobilisomes were incubated for 16 h at 37 °C with 70% formic acid in before adding SDS sample buffer and analysis by Tricine SDS-PAGE on a 17% polyacrylamide gel.

#### **Isolation of Protein Extracts from *Synechocystis* sp. PCC 6803**

*Synechocystis* sp. PCC 6803 cells were grown and harvested as described above. The cell pellet was resuspended in TEN100 buffer [42], containing PIC. Cell lysis was performed as described above.

#### **MALDI-TOF MS Analysis**

The protein spots were subjected to in-gel trypsin digestion before mass spectrometry analysis as described previously [43]. The peptide mixtures from the tryptic digests were desalted and concentrated using ZipTips™ columns made from the reverse chromatography resins Poros and Oligo R3 (Applied Biosystems). The bound peptides were washed with a solution of 0.5% formic acid and eluted from the column in 1 µl of 33% (v/v) acetonitrile/0.1% trifluoroacetic acid solution saturated with α-cyano-4-hydroxycinnamic (Bruker Daltonics) directly onto a MALDI target plate and air dried before analysis in the mass spectrometer. Mass spectrometry measurement was performed on an Ultraflex-TOF TOF tandem mass spectrometer (Bruker Daltonics). Peptide mass fingerprint spectra were acquired in the reflectron positive mode with a pulsed extraction using approximately 100 laser shots. The spectra were acquired after an external calibration using reference peptides (Peptide mixture II Bruker Daltonics). The acquired spectra were further internally calibrated using trypsin autolysis peaks as internal standards (842.5100, 2211.1046 Da). Monoisotopic masses were assigned and processed using Biotoools™ and FlexAnalysis™ software (Bruker Daltonics) before submitting them to the Mascot program [44] for searches against the non-redundant NCBI database. The parameters used in the Mascot peptide mass fingerprint searches were as follows: Taxonomy, *Synechocystis*; search all molecular masses and all isoelectric points; allow up to one missed proteolytic cleavage site and a peptide mass tolerance of 50 ppm. Methionine oxidation was considered as an optional modification and cysteine carbamidomethylation as a fixed modification in all the searches. Matches to *Synechocystis* proteins were considered unambiguous when the probability score was significant using the Mascot score with a p value < 0.05 and when there was a minimum of five peptides and with a sequence coverage greater than 20%. For each protein the identity was further validated by tandem MS-MS analysis of selected peptides.

#### **In silico Analysis**

Blast search analyses were done by NCBI protein-protein blast [45] (see Additional file 1). This program was also used for conserved domain predictions. Transmembrane domains were predicted by using TMHMM server v. 2.0 [46] and the SOSUI protein prediction server [47]. Genome data from *Synechocystis* sp. PCC 6803 were obtained from CyanoBase [48]. *In silico* cleavage was performed by PeptideMass [49].

#### **Network Construction**

The 41 sequences were aligned using MUSCLE [50] with 16 iterations. The output format was set to the standard ClustalW format [51]. The alignment contained 307 sites including 191 gapped sites that were excluded from the

analysis. From the remaining 116 sites logdet distances were estimated using LDDist 1.3.2 [52]. These distances were used to construct a Neighbor-net splits graph [53] that was visualized with SplitsTree4 [54].

### Authors' contributions

KB, OK and JP performed the experiments except phylogenetic NG and MALDI-TOF analyses JN, UGM initiated the project and was the supervisor. The manuscript, to which all authors contributed, was designed and written by KB, OK and UGM All authors have read and approved of the final version of the manuscript.

### Additional material

#### Additional file 1

Accession numbers. Sequences were obtained from NCBI and the Cyanidioschyzon merolae Genome Project with the exception of *cpcT* from *Synechococcus sp. PCC 7002*, which was downloaded from [10].

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2229-8-56-S1.pdf>]

### Acknowledgements

We thank Andrew Bozarth for comments on the manuscript. This work is supported by the Deutsche Forschungsgemeinschaft (SFB-TR1, TP A7).

### References

- Bishop JE, Rapoport H, Klotz CV, Chan CF, Glazer AN, Fuglistaller P, Zuber H: **Chromopeptides from phycoerythrocyanin: structure and linkage of the three bilin groups.** *J Am Chem Soc* 1987, **109**:875-881.
- Stadnichuk I: **Phycobiliproteins: Determination of chromophore composition and content.** *Phytochemical Analysis* 1995, **6**:281-288.
- Holzwarth AR: **Structure function relationships and energy transfer in phycobiliprotein antennae.** *Physiol Plant* 1991, **83**:518-528.
- Fairchild CD, Zhao J, Zhou J, Colson SE, Bryant DA, Glazer AN: **Phycocyanin alpha-subunit phycocyanobilin lyase.** *Proc Natl Acad Sci USA* 1992, **89**:7017-7021.
- Fairchild CD, Glazer AN: **Oligomeric structure, enzyme kinetics, and substrate specificity of the phycocyanin alpha subunit phycocyanobilin lyase.** *J Biol Chem* 1994, **269**:8686-8694.
- Jung LJ, Chan CF, Glazer AN: **Candidate genes for the phycoerythrocyanin alpha subunit lyase. Biochemical analysis of *pecE* and *pecF* interposon mutants.** *J Biol Chem* 1995, **270**:12877-12884.
- Zhao KH, Deng MG, Zheng M, Zhou M, Parbel A, Storf M, Meyer M, Strohmman B, Scheer H: **Novel activity of a phycobiliprotein lyase: both the attachment of phycocyanobilin and the isomerization to phycoviolobilin are catalyzed by the proteins *PecE* and *PecF* encoded by the phycoerythrocyanin operon.** *FEBS Lett* 2000, **469**:9-13.
- Storf M, Parbel A, Meyer M, Strohmman B, Scheer H, Deng MG, Zheng M, Zhou M, Zhao KH: **Chromophore attachment to biliproteins: specificity of *PecE/PecF*, a lyase-isomerase for the photoactive 3(1)-cys-alpha 84-phycoviolobilin chromophore of phycoerythrocyanin.** *Biochemistry* 2001, **40**:12444-12456.
- Zhao KH, Su P, Li J, Tu JM, Zhou M, Bubenzer C, Scheer H: **Chromophore attachment to phycobiliprotein beta-subunits: phycocyanobilin:cysteine-beta84 phycobiliprotein lyase activity of *CpeS*-like protein from *Anabaena Sp. PCC7120*.** *J Biol Chem* 2006, **281**:8573-8581.
- Shen G, Saunee NA, Williams SR, Gallo EF, Schluchter WM, Bryant DA: **Identification and characterization of a new class of bilin lyase: the *cpcT* gene encodes a bilin lyase responsible for attachment of phycocyanobilin to Cys-153 on the beta-subunit of phycocyanin in *Synechococcus sp. PCC 7002*.** *J Biol Chem* 2006, **281**:17768-17778.
- Zhao KH, Su P, Tu JM, Wang X, Liu H, Ploscher M, Eichacker L, Yang B, Zhou M, Scheer H: **Phycobilin:cysteine-84 biliprotein lyase, a near-universal lyase for cysteine-84-binding sites in cyanobacterial phycobiliproteins.** *Proc Natl Acad Sci USA* 2007, **104**:14300-14305.
- Zhao KH, Zhang J, Tu JM, Bohm S, Ploscher M, Eichacker L, Bubenzer C, Scheer H, Wang X, Zhou M: **Lyase activities of *CpcS*- and *CpcT*-like proteins from *Nostoc PCC7120* and sequential reconstitution of binding sites of phycoerythrocyanin and phycocyanin beta-subunits.** *J Biol Chem* 2007, **282**:34093-34103.
- Glazer AN: **Light guides. Directional energy transfer in a photosynthetic antenna.** *J Biol Chem* 1989, **264**:1-4.
- Apt KE, Collier JL, Grossman AR: **Evolution of the phycobiliproteins.** *J Mol Biol* 1995, **248**:79-96.
- MacColl R: **Cyanobacterial phycobilisomes.** *J Struct Biol* 1998, **124**:311-334.
- Bryant DA, Glazer AN, Eiserling FA: **Characterization and structural properties of the major biliproteins of *Anabaena sp.*** *Arch Microbiol* 1976, **110**:61-75.
- Glazer AN: **Phycobiliproteins.** *Methods Enzymol* 1988, **167**:291-303.
- de Marsac NT, Cohen-Bazire G: **Molecular composition of cyanobacterial phycobilisomes.** *Proc Natl Acad Sci USA* 1977, **74**:1635-1639.
- Lundell DJ, Williams RC, Glazer AN: **Molecular architecture of a light-harvesting antenna. In vitro assembly of the rod substructures of *Synechococcus 6301* phycobilisomes.** *J Biol Chem* 1981, **256**:3580-3592.
- Liu LN, Chen XL, Zhang YZ, Zhou BC: **Characterization, structure and function of linker polypeptides in phycobilisomes of cyanobacteria and red algae: an overview.** *Biochim Biophys Acta* 2005, **1708**:133-142.
- Glazer AN: **Light harvesting by phycobilisomes.** *Annu Rev Biophys Biophys Chem* 1985, **14**:47-77.
- Douglas SE, Penny SL: **The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae.** *J Mol Evol* 1999, **48**:236-244.
- Gould SB, Fan E, Hempel F, Maier UG, Klosgen RB: **Translocation of a phycoerythrin alpha subunit across five biological membranes.** *J Biol Chem* 2007, **282**:30295-30302.
- Matsuzaki M, Misumi O, Shin IT, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Yoshida Y, Nishimura Y, Nakao S, Kobayashi T, Momoyama Y, Higashiyama T, Minoda A, Sano M, Nomoto H, Oishi K, Hayashi H, Ohta F, Nishizaka S, Haga S, Miura S, Morishita T, Kabeya Y, Terasawa K, Suzuki Y, Ishii Y, Asakawa S, Takano H, Ohta N, Kuroiwa H, Tanaka K, Shimizu N, Sugano S, Sato N, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T: **Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae 10D*.** *Nature* 2004, **428**:653-657.
- Asano T, Yoshioka Y, Kurei S, Sakamoto W, Machida Y: **A mutation of the *CRUMPLED LEAF* gene that encodes a protein localized in the outer envelope membrane of plastids affects the pattern of cell division, cell differentiation, and plastid division in *Arabidopsis*.** *Plant J* 2004, **38**:448-459.
- Mann NH, Clokie MR, Millard A, Cook A, Wilson WH, Wheatley PJ, Letarov A, Krisch HM: **The genome of *S-PM2*, a "photosynthetic" T4-type bacteriophage that infects marine *Synechococcus* strains.** *J Bacteriol* 2005, **187**:3188-3200.
- Cavalier-Smith T: **Nucleomorphs: enslaved algal nuclei.** *Curr Opin Microbiol* 2002, **5**:612-619.
- Maier UG, Douglas SE, Cavalier-Smith T: **The nucleomorph genomes of cryptophytes and chlorarachniophytes.** *Protist* 2000, **151**:103-109.
- Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reich M, Cavalier-Smith T, Maier UG: **The highly reduced genome of an enslaved algal nucleus.** *Nature* 2001, **410**:1091-1096.
- Federspiel NA, Grossman AR: **Characterization of the light-regulated operon encoding the phycoerythrin-associated linker**

- proteins from the cyanobacterium *Fremyella diplosiphon*. *J Bacteriol* 1990, **172**:4072-4081.
31. Federspiel NA, Scott L: **Characterization of a light-regulated gene encoding a new phycoerythrin-associated linker protein from the cyanobacterium *Fremyella diplosiphon***. *J Bacteriol* 1992, **174**:5994-5998.
  32. de Lorimier R, Bryant DA, Stevens SE Jr: **Genetic analysis of a 9 kDa phycocyanin-associated linker polypeptide**. *Biochim Biophys Acta* 1990, **1019**:29-41.
  33. Ughy B, Ajlani G: **Phycobilisome rod mutants in *Synechocystis* sp. strain PCC6803**. *Microbiology* 2004, **150**:4147-4156.
  34. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY: **Generic assignments, strain histories and properties of pure cultures of cyanobacteria**. *J Gen Microbiol* 1979, **111**:1-61.
  35. Gray BH, Lipschultz CA, Gantt E: **Phycobilisomes from a blue-green alga *Nostoc* species**. *J Bacteriol* 1973, **116**:471-478.
  36. Laemmli UK: **Cleavage of structural proteins during the assembly of the head of bacteriophage T4**. *Nature* 1970, **227**:680-685.
  37. Schagger H, von Jagow G: **Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa**. *Anal Biochem* 1987, **166**:368-379.
  38. Berkelman TR, Lagarias JC: **Visualizalation of bilin-linked peptides and proteins in polyacrylamide gels**. *Anal Biochem* 1986, **156**:194-201.
  39. Zolla L, Bianchetti M: **High-performance liquid chromatography coupled on-line with electrospray ionization mass spectrometry for the simultaneous separation and identification of the *Synechocystis* PCC 6803 phycobilisome proteins**. *J Chromatogr A* 2001, **912**:269-279.
  40. Wessel D, Flugge UI: **method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids**. *Anal Biochem* 1984, **138**:141-143.
  41. Piszkiwicz D, Landon M, Smith EL: **Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations**. *Biochem Biophys Res Commun* 1970, **40**:1173-1178.
  42. Frangioni JV, Neel BG: **Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins**. *Anal Biochem* 1993, **210**:179-187.
  43. Hellmann U, Wernstedt C, Gonez J, Heldin CH: **Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing**. *Anal Biochem* 1995, **224**:451-455.
  44. Mascot Program [<http://www.matrixscience.com>]
  45. NCBI protein-protein blast [<http://www.ncbi.nlm.nih.gov/BLAST>]
  46. TMHMM server v. 2.0 [<http://www.cbs.dtu.dk/services/TMHMM-2.0/>]
  47. SOSUI protein prediction server [<http://bp.nuap.nagoya-u.ac.jp/sosui/>]
  48. CyanoBase [<http://bacteria.kazusa.or.jp/cyanobase>]
  49. PeptideMass [<http://www.expasy.ch/tools/peptide-mass.html>]
  50. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucleic Acids Res* 2004, **32**:1792-1797.
  51. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**. *Nucleic Acids Res* 1994, **22**:4673-4680.
  52. Thollesson M: **LDDist: a Perl module for calculating LogDet pair-wise distances for protein and nucleotide sequences**. *Bioinformatics* 2004, **20**:416-418.
  53. Bryant D, Moulton V: **Neighbor-net: an agglomerative method for the construction of phylogenetic networks**. *Mol Biol Evol* 2004, **21**:255-265.
  54. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies**. *Mol Biol Evol* 2006, **23**:254-267.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

