Structural and Functional Characterization of PII and PII-like Proteins and their Network of Interactions

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Khaled Ali Mohamed Ali Selim
aus Kairo

Tübingen
2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	24.04.2019
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Karl Forchhammer
2. Berichterstatter:	Prof. Dr. Klaus Hantke
3. Berichterstatter:	Prof. Dr. Gottfried Unden

Erklärung

Hiermit erkläre ich, dass ich die zur Promotion eingereichte Arbeit selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und Stellen, die wörtlich oder inhaltlich nach den Werken anderer Autoren entnommen sind, als solche gekennzeichnet habe. Eine detaillierte Abgrenzung meiner eigenen Leistungen von den Beiträgen meiner Kooperationspartner habe ich in „List of publications and declaration of personal contribution“ vorgenommen.

Tübingen, den tt,mm,jjjj ----------

Life is a journey . . . from birth to death. Countries change, people change, even we change, and our thoughts change too, but life keeps going and we have to put up with it!

سَنُرِيهِمْ آيَاتِنَا فِي الْآفَاقِ وَفِي أَنفُسِهِمْ حَتَّىٰ يَتَبَيَّنَ لَهُمْ أَنَّهُ الْحَقُّ ۗ أَوَلَمْ يَكْفِ بِرَبِّكَ أَنَّهُ عَلَىٰ كُلِّ شَيْءٍ شَهِيدٌ (53)

سورة فصلت - القرآن الكريم

„Wir werden sie in der weiten Welt und in ihnen selber unsere Zeichen sehen lassen damit ihnen klar wird, dass es die Wahrheit ist. Genügt es denn nicht, dass dein Herr über alles Zeuge ist? “

Der edle Qur'an [Sure 41: Vers 53]

Table of contents

Acknowledgments

First and foremost, my great praise, and sincere thanks should be submitted to almighty GOD, the most gracious, and the most merciful, for the continuous support to me, surpassing all the obstacles I faced, and gave us the knowledge and the ability to think and believe. By the grace of ALLAH, this research work has been accomplished.

With the deepest gratitude I wish to thank every person who has come into my life and inspired, touched and illuminated me through his or her presence. Also, I would like to acknowledge and express my gratitude to the following people for their magnificent support and contributions to the creation of this research work:

- The words fall to express my deep thanks, appreciation and sincere gratitude to Prof. Dr. Karl Forchhammer for his close supervision, guidance, indispensable valuable advices, and the continuous support throughout this research work. He always had the time for questions and discussions. I find the German word “Doktorvater” suits him very good, since I truly consider him my father in science.
- I would like to express my deep thanks to Prof. Dr. Klaus Hantke for his interest in this research work, for his support, and for his time and efforts to referee my dissertation.
- I am deeply indebted to Dr. Marcus D. Hartmann for his kind assistance, the stimulating debates and suggestions to improve my thesis, and for encouraging me in the research and the writing of this thesis. I owe a great deal to him for teaching me the basics of crystallography, accuracy, and patience.
- I would like to deliver special thanks to PD. Dr. Iris Maldener for being always there when I needed help and advices.
- I would like to thank also Prof. Dr. Andrei Lupas for giving me the opportunity to work within MPI society, for his interest in this research work, and for his support with MPI-tools.
- I also appreciate the time and efforts of Prof. Dr. Friedrich Götz (Chairman of my Ph.D. defense committee) and PD. Dr. Iris Maldener for examining the thesis during the Ph.D. defense.
- I would like to express my deepest thanks, gratitude and appreciation to Dr. Jan Lüddecke and Dr. Jan Bornikoel for sparing no effort or time to help me during my first months in the Lab, teaching me several techniques, and for the precious suggestions.

- I would like to express my sincere gratitude to Jörg Scholl for proofreading my thesis and helping me with the German Abstract.
- I am also indebted by deep thanks to my current and former lab mates of ''Raum-9A08'', Dr. Rebeca Pérez, Eva Bok, Dmitry Shvarev, Dr. Carolina Nishi, and Ritu Taigwal, for the comfortable and friendly atmosphere, valuable and inspiring discussions, and continuous encouragement.
- Sincere appreciation and deep thanks to my students Michael Haffner, Larissa Dietz, and Lingxi Gao.
- Special thanks are owed to Sofia Doello and Carolin Sokele for the funny times after stressful working days, volleyball training, and Ice-cream breaks.
- I would like to acknowledge DAAD (German Academic Exchange Service) for supporting me with a scholarship, which gave me the opportunity to expose to new culture and helped me to develop and expand my knowledge and scientific experience.
- Last, but not least, my thanks to all members and colleagues of AG Forchhammer for their kind dealings, sincere cooperation, and encouragement during execution of this research work. You really made the time in Tübingen to be so much enjoyable!
- Finally, I cannot find the words that express my appreciation, grateful and heartfelt gratitude to my family. From the deep of my heart, I would like to thank Dr. Ali Selim, my lovely mother, my Nina, my brothers Ahmed, AbdEl-Rahman and Mómen, for their everlasting support in all aspects of life, unlimited and continuous encouragement, true love, care and strong motivation for finishing this work.

Dedicated

- To the souls of my uncle (El Arabi) and my grandmother (Um-El Arabi)
- To my parents
- To Nina

For their inspiration, continuous love, unlimited encouragement and continuous support.

Abbreviations

µE: Micro-Einstein
2-OG: 2-oxoglutarate
2-PG: 2-phosphoglycolate
3-PGA: 3-phosphoglycerate
ACCase: Acetyl-CoA carboxylase
ADP: adenosine diphosphate
AMP: adenosine monophosphate
AmtB: Ammonium transport channel
AR: adenylyl-removing
ATase: adenylyl-transferase
ATP: adenosine triphosphate
BC: Biotin carboxylase subunit of ACCase
BCCP: Biotin carboxyl carrier protein subunit of ACCase
CA: Carbonic anhydrase
Ca^{2+}: Calcium cation
cAMP: Second messenger cyclic-AMP
CCM: Carbon concentrating mechanisms
c-di-AMP: Second messenger cyclic-di-AMP
C_i: Inorganic carbon
CLANS: CLuster ANalysis of Sequences program
CO_2: Carbon dioxide
CoA: Coenzyme A
CSE: Ca^{2+} sensor EF-hand protein
CT: Carboxyltransferase subunit of ACCase
CutA: divalent tolerance protein
cya1: Gene encoding for adenylyl cyclase
CyanoBase: Data-base for cyanobacteria genomes
Gln: Glutamine
glnB and *glnK*: Genes encoding for canonical PII protein
Glu: Glutamate
GOGAT: Glutamate synthetase (Glutamine oxoglutarate aminotransferase)
GS: Glutamine synthetase
HCO_3^-: Bicarbonate
ITC: Isothermal titration calorimetry
MALS: Multi angle light scattering
N_2: Nitrogen
NAGK: N-acetyl-L-glutamate kinase
NH_4^-: Ammonium
O_2: Oxygen
OD: Optical density
PAM: Pulse amplitude modulation
PEP: Phosphoenolpyruvate
PII-NG: PII new group (refer to PII in associate with heavy metal efflux pump; HmeP)
PipX: PII-interacting partner X
PSII: Photo system II
RuBisCO: CO_2 fixing enzyme (Ribulose-1,5-bisphosphate carboxylase/oxygenase)
sAC: Soluble adenylyl cyclase
SbtA: Sodium-dependent HCO_3^- transporter
SbtB: 2nd gene of SbtA operon
SEC: Size exclusion chromatography
SPR: Surface plasmon resonance
TCA: tricarboxylic acid
WT: Wild type

A. Abstract

The signaling transduction proteins of PII superfamily are one of the ancient and highly conserved protein families in nature and widely distributed in all domains of life. Canonical PII exerts it's signaling function via the binding of small effector molecules, including ATP, ADP, and 2-oxoglutarate (2-OG). All cyanobacteria contain a PII homologue of the GlnB-type, which is involved in the regulation of nitrogen assimilation by binding to N-acetyl-L-glutamate kinase (NAGK), the key enzyme of arginine biosynthesis, and to PII-interacting protein X (PipX), the co-activator of the global nitrogen-trascription factor NtcA. The binding of PII to NAGK enhances NAGK activity and prevents arginine (Arg) feedback inhibition, thereby increasing the flux into the arginine pathway. Formation of the PII-PipX complex causes a deactivation of the NtcA regulon. In this study, we identified Lys58 to be a key residue for proper sensing/signaling function of cyanobacterial PII protein. Furthermore, this study provided the first biochemical characterization of red algal PII protein from *Porphyra purpurea.* Its properties represent an intermediary state between PII properties in Cyanobacteria and Chlorophyta. A comparison to Chlorophyta PII proteins showed that during later stages of evolution of Chlorophyta, the PII proteins diverged in their properties, becoming very heterogeneous with respect to ADP and to 2-OG binding, while only the binding of ATP in presence of 2-OG has been conserved. Further, we characterized the first metabolic adaptation strategies of colorless-nonphotosynthetic alga *Polytomella parva*, in comparison to closely related green-photosynthetic alga *Chlamydomonas reinhardtii*, in response to nitrogen starvation. Metabolome analysis revealed that *P. parva* accumulates higher amounts of TCA intermediates as well as arginine, glutamate and aspartate. Strikingly, the PII and NAGK proteins in *P. parva* have coevolved into a stable hetero-oligomeric complex, irrespective of effector molecules such that PII evolved into a subunit of NAGK.

The PII-like proteins lack the canonical PII signature sequences but are structurally clearly related to classical PII proteins. Their functions, targets and regulatory responses are unknown. A close examination of available cyanobacteria genomes revealed several different PII-like proteins. In this study, we focused on two PII-like proteins in cyanobacteria namely: 1) SbtB protein, which is located in one operon next to sodium dependent bicarbonate transporter (SbtA) and 2) the divalent ion tolerance protein CutA. Biochemical and physiological studies showed that the SbtB protein is an important component of the cyanobacterial carbon concentration mechanism (CCM) for sensing the fluctuation levels of inorganic carbon (C_i) via binding to the secondary messenger cAMP. Also, we demonstrated that cyanobacterial CutA protein is not involved directly in heavy metal sensing. SbtB was revealed as the first protein of the PII superfamily, which specifically

binds cAMP in addition to ATP, ADP and AMP. However, for CutA, no binding of effector molecules could be identified. I resolved the crystal structures of CutA (in apo form) and SbtB (in apo status and in complex with AMP and cAMP) proteins. The nucleotide-binding pocket of SbtB is located between the lateral cleft of two subunits, as in canonical PII proteins. Apparently, the trimeric core architectural principle of PII-like proteins of CutA and SbtB are widely distributed similar to the canonical PII core architecture. This clearly indicates that the proteins of the PII superfamily arose evolutionary from a common trimeric ancestor protein.

B. Zusammenfassung

Die Signaltransduktionsproteine der PII-Superfamilie stellen eine der ältesten hochkonservierten Proteinfamilien dar und finden sich in allen Domänen des Lebens. Kanonische PII-Signalproteine üben ihre Funktion dadurch aus, dass sie als Effektormoleküle ATP, ADP und 2-Oxoglutarat (2-OG) binden. Alle bislang untersuchten Cyanobakterien enthalten PII-Homologe des GlnB-Typs. Es reguliert den zellulären Stickstoff-Stoffwechsel, indem es an die N-Acetyl-L-Glutamat-Kinase (NAGK, katalysiert den geschwindigkeitsbestimmenden Schritt der Arginin-Biosynthese) sowie an das PII-Interacting Protein X (PipX, Co-Aktivator des globalen Stickstoff-Transkriptionsfaktors NtcA) bindet. Die Interaktion von PII mit NAGK führt zur Erhöhung der NAGK-Aktivität und unterbindet die hemmende Wirkung von Arginin auf NAGK, was letztlich zu einer verstärkten Produktion von Arginin führt. Die Bildung eines PII-PipX-Komplexes wiederum führt zur Deaktivierung des NtcA-Regulons. In dieser Arbeit konnte gezeigt werden, dass der Lys58-Rest eine Schlüsselrolle bei der Signaltransduktion des cyanobakteriellen PII spielt. Das PII-Protein aus der Rotalge *Porphyra purpurea* wurde erstmals biochemisch charakterisiert. Es stellt einen Übergangszustand zwischen den PII-Proteinen aus Cyanobakterien und Chlorophyta dar. Durch den Vergleich der sensorischen Eigenschaften des PII-Proteins aus Rotalgen mit PII-Proteinen aus verschiedenen Phyla oxygener phototropher Organismen (Cyanobakterien, Grünalgen und Moose) zeigte sich, dass die PII-Signaltransduktion in Rotalgen weitgehend der in Cyanobakterien gleicht, wohingegen sich die Eigenschaften der PII-Proteine im Laufe der Entwicklung der Chlorophyta verändert haben. So verhalten sie sich sehr heterogen bezüglich der Bindung von ADP und 2-OG, nur die Bindung von ATP in Gegenwart von 2-OG ist konserviert. Eine weitere Studie befasste sich mit der PII-Signaltransduktion in der nicht photosynthetisch aktiven Alge *Polytomella parva* im Vergleich zur nahe verwandten, Photosynthese betreibenden Alge *Chlamydomonas reinhardtii.* Metabolom-Daten zufolge akkumuliert *P. parva* TCA-Intermediate sowie Arginin, Glutamat und Aspartat. Ferner zeigte diese Studie dass sich im Lauf der *P. parva* Evolution PII und NAGK zu einem stabilen, heterooligomeren Komplex entwickelt haben hat, der nicht mehr durch Effektormoleküle dissoziiert wird. Dementsprechend kann PII in diesem Fall als Untereinheit von NAGK angesehen werden.

PII-ähnliche Proteine ähneln strukturell den klassischen PII-Proteinen, besitzen aber nicht die kanonischen PII Signaturmotive. Bislang war nichts über ihre Funktionen, Interaktionspartner und regulatorischen Reaktionen bekannt. Eine *in-silico* Analyse der Genome von Cyanobakterien zeigte das Vorhandensein verschiedener PII-ähnlicher Proteine (PII-like proteins). Im Rahmen dieser Arbeit lag der Fokus auf zwei PII-ähnlichen Proteinen in Cyanobakterien: 1) SbtB, dessen

Gen in einem Operon in direkter Nachbarschaft zum Natriumabhängigen Bicarbonat-Transporter (SbtA) liegt und 2) CutA (divalent ion tolerance protein). Es konnte sowohl biochemisch als auch physiologisch gezeigt werden, dass SbtB eine wichtige Komponente des Kohlenstoffanreichungsmechanismus (carbon concentration mechanism, CCM) ist, indem es Änderungen der intrazellulären Konzentration von anorganischem Kohlenstoff (C_i) durch Bindung des sekundären Messengers cAMP verarbeitet. Darüber hinaus konnten wir zeigen, dass CutA in Cyanobakterien nicht direkt an der Erkennung von Schwermetallionen beteiligt ist. Weiterhin konnten wir für SbtB als erstes Mitglied der PII-Superfamilie zeigen, dass neben den bekannten Effektoren ATP, ADP und AMP auch cAMP spezifisch gebunden wird. Jedoch blieben die sensorischen Eigenschaften von CutA unklar. Die Kristallstrukturen von CutA (Apo-Form) und SbtB (Apo-Form und im Komplex mit AMP und cAMP) konnte aufgeklärt werden. Dadurch konnte die Nukleotid-Bindetasche von SbtB in der lateralen Falte zweier Untereinheiten lokalisiert werden. Offenbar ist die trimere Grundstruktur der PII-ähnlichen Proteine SbtB und CutA ähnlich wie die der kanonischen PII-Proteine weit verbreitet. Dies zeigt eindeutig, dass die Proteine der PII-Superfamilie evolutionär aus einem gemeinsamen Vorfahren hervorgegangen sind.

C. List of publications and declaration of personal contribution

1. Accepted publications

1.1. Publication 1 (Research Article):

Lapina T[#], **Selim KA**[#], Forchhammer K, Ermilova E. (2018) The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins. Sci Rep. 8(1):790.

[#]Share first co-author and contributed equally to this work.

1.2. Publication 2 (Research Article):

Selim KA, Haase F, Hartmann MD, Hagemann M, Forchhammer K. (2018) PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. Proc Natl Acad Sci U S A. 115(21): E4861-E4869.

1.3. Publication 3 (Research Article):

Selim KA, Lapina T, Forchhammer K, Ermilova E. (2019) Interaction of N-acetyl-l-glutamate kinase with the PII signal transducer in the non-photosynthetic alga *Polytomella parva*: Co-evolution towards a hetero-oligomeric enzyme. FEBS J. doi: 10.1111/febs.14989.

1.4. Publication 4 (Research Article):

Walter J, **Selim KA**, Leganés F, Fernández-Piñas F, Vothknecht UC, Forchhammer K, Aro E-M, Gollan PJ. (2019) A novel Ca^{2+}-binding protein influences photosynthetic electron transport in *Anabaena* sp. PCC 7120. Biochim Biophys Acta Bioenerg. 1860(6):519-532.

1.5. Publication 5 (Research Article):

Selim KA, Haffner M, Watzer B, Forchhammer K. (2019) Tuning the *in vitro* sensing and signaling properties of cyanobacterial PII protein by mutation of key residues. Sci Rep. 9(1): 18985.

2. Submitted manuscript

2.1. Manuscript 1 (Research Article):

Selim KA, Alva V, Hartmann MD, Forchhammer K. Structural and functional characterization of cyanobacterial PII-like protein CutA does not hint at an involvement in heavy metal tolerance. FEBS J. (Submitted 2019).

3. Declaration of personal contribution to the publications

Publication 1: *The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins.*

I overexpressed and purified cyanobacteria proteins in this work. I planned, preformed, analyzed and interpreted all ITC experiments in this research article to characterize the relevant sensory properties of PII protein from the red alga *Porphyra purpurea* in comparison to PII proteins from different phyla of oxygenic phototrophs (cyanobacteria, *Chlamydomonas* and *Physcomitrella*). Furthermore, I participated in the writing of the manuscript with inputs from Prof. Dr. Karl Forchhammer and Prof. Dr. Elena Ermilova. I prepared the tables and figures, submitted the manuscript to the journal, and responded to the reviewers comments under the supervision of Prof. Dr. Karl Forchhammer.

Publication 2: *PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response.*

All molecular biology work including creation of: 1) the knockout mutants of Δ*sbtB* and Δ*cya1* in cyanobacterium *Synechocystis* sp. PCC 6803, 2) complemented GFP-expressing strains, 3) construct for overexpression of recombinant SbtB protein, was done by me. I planned, preformed, analyzed and interpreted all biochemical/biophysical experiments in this work to characterize *in vitro* and *in vivo* properties of SbtB protein (protein overexertion and purification, size exclusion chromatography, ITC, MST, immuno-blotting analysis, PAM-measurements, and O_2 evolution). With the help of Dr. Marcus D. Hartmann, I solved the crystal structures of SbtB protein to unravel the binding modes of SbtB with AMP and the secondary massager cAMP and submitted the atomic coordinates to protein data bank. The microscopy was performed by me. I generated a part of the physiological data including: some of growth and O_2 evolution experiments for the Δ*sbtB* and the complemented mutants, estimating the fluorescent levels of SbtB-GFP protein for the complemented strains, and the PAM analysis (Figs. 2*B & C* and S7). Furthermore, I analyzed

and interpreted all the generated data, wrote the manuscript with inputs from all co-authors, and prepared the tables and figures. I am the corresponding author for this research article; therefore, I submitted the manuscript to the journal and responded to the reviewers comments under the supervision of Prof. Dr. Karl Forchhammer.

Publication 3: *Interaction of N-acetyl-l-glutamate kinase with the PII signal transducer in the non-photosynthetic alga Polytomella parva: Co-evolution towards a hetero-oligomeric enzyme.*
I and Tatiana Lapina contributed equally in the generation of the experimental data for this manuscript. I planned and preformed: 1) metabolomic analysis (the samples were submitted to analytics unite, ZMBP-Tübingen University for LC/MS-MS analysis), 2) some of analytical size exclusion chromatography coupled to multi angles light scattering (SEC-MALS) runs, 3) surface plasmon resonance (SPR) spectroscopy. Furthermore, I analyzed and interpreted all the generated data, wrote the manuscript under the supervision of Prof. Dr. Karl Forchhammer and with inputs from Prof. Dr. Elena Ermilova, and prepared the tables and figures. I submitted the manuscript to the journal and responded to the reviewers comments under the supervision of Prof. Dr. Karl Forchhammer.

Publication 4: *A novel Ca^{2+}-binding protein influences photosynthetic electron transport in Anabaena sp. PCC 7120.*
Contribution: I planned, performed and analyzed the biochemical characterization of the recombinant calcium-binding protein, including ITC, SEC-MALS and CD-spectra. I wrote and prepared the figures and the table of the corresponding sections (results and material & methods) for the manuscript.

Publication 5: *Tuning the in vitro sensing and signaling properties of cyanobacterial PII protein by mutation of key residues.*
All the molecular biology work for the creation of overexpression constructs for recombinant point-mutation variants of PII protein from *Synechococcus elongatus* PCC 7942, was done by me (except for I86N variant of PII). All the biochemical experiments for characterization PII variants in this manuscript were designed and performed by me or under my supervision, except for the SPR analysis for PII (I86N) variant. I analyzed and interpreted all the generated data (except for I86N variant of PII), wrote the manuscript under the supervision of Prof. Dr. Karl Forchhammer, and prepared the tables and figures except for the SPR figure for PII (I86N) variant. I am the

corresponding author of this research article; therefore, I submitted the manuscript to the journal and responded to the reviewers comments.

Manuscript 1: *Structural and functional characterization of cyanobacterial PII-like protein CutA does not hint at an involvement in heavy metal tolerance.*

I performed the search in protein data bank to identify proteins with PII-ferredoxin-like fold and fed the sequences to Vikram Alva, who did the enrichment and the clustering of the sequences. I generated all the CutA mutants listed in this manuscript and created constructs for overexpression and purification of CutA proteins. All the experimental data in this manuscript was generated by me, except for the bioinformatic analysis. I solved the crystal structure of CutA protein from cyanobacterium *Nostoc* sp. PCC7120 under the supervision of Dr. Marcus D. Hartmann. I analyzed and interpreted all the generated data, wrote the manuscript under the supervision of Prof. Dr. Karl Forchhammer, with inputs from all the co-authors, and prepared the tables and figures.

D. Introduction

1. Oxygenic photoautotrophs

Aerobic life on Earth evolved about 3.8-2.7 billion years ago with the evolution of oxygenic photosynthesis by cyanobacteria (Hayes 1996, Mojzsis et al. 1996, Hohmann-Marriott & Blankenship 2011). Approximately 2.4 billion years ago, the ancestors of today's cyanobacteria were the first oxygenic photoautotrophs to release molecular O_2, as a waste product of the oxygenic photosynthesis (Bekker et al. 2004, Kerr 2005, Holland 2006, Lyons et al. 2014). Via endosymbiosis, the photosynthetic ability was transmitted later from cyanobacteria to eukaryotes, giving rise to plastids in algae and plants (Hohmann-Marriott & Blankenship 2011). Both, cyanobacterial and plant cells process the photosystems I and II (Blankenship & Hartman 1998). The high structural and functional similarity of the cyanobacterial photosynthetic apparatus with higher plants can be explained by the primary endosymbiosis that happened about 1.8 billion years ago (Gould et al. 2008, Archibald 2009, Hohmann-Marriott & Blankenship 2011, Tirichine & Bowler 2011). In the course of evolution, the cyanobacterial cells were integrated within the eukaryotic cells to form plastids and chloroplasts (i.e. eukaryotes photosynthetic organelles).

The photoautotrophic life style allows the conversion of inorganic carbon (C_i) sources such as carbon dioxide (CO_2) and bicarbonate (HCO_3^-) into organic carbon at the expense of light energy. As a consequence of light reactions, the atmospheric O_2 accumulated due to the utilization of water as an electron donor. Then, the electron is transferred through the photosynthetic complexes to reduce $NADP^+$ (Binder 1982). Concomitantly with CO_2 fixation into organic compounds to build up carbohydrates for metabolic activities or for storage, the accumulation of atmospheric O_2 due to the cyanobacterial water-splitting activity suited the earth's atmosphere for evolution of complex life forms. Overtime, the activity of photoautotrophic organisms led to a drop in the amount of available C_i on Earth, especially in aquatic habitats. To adapt to the drop of C_i availability, cyanobacteria and later algae evolved the carbon concentrating mechanism (CCM) that is essential to grow in the current Earth's atmosphere containing about 0.04 % CO_2. In general, cyanobacterial CCM is composed of different C_i uptake systems and the carboxysome, a protein-coated bacterial cellular compartment where the CO_2 fixing enzyme RubisCO and carbonic anhydrase (CA) are localized. The C_i uptake systems massively concentrate HCO_3^- inside the cyanobacterial cells, which is then diffusing into the carboxysome where CA converts HCO_3^- into CO_2. This strategy increases the local concentrations of the CO_2 nearby the RubisCO to minimize the oxygenation activity and enhance the carboxylation reaction (Rae et al. 2013, Burnap et al. 2015).

Apart from the evolutionary-historical role of cyanobacteria in the invention of oxygenic photosynthesis for development of aerobic metabolism on Earth, cyanobacteria are among the main primary producers. They are responsible for a high proportion of CO_2 fixation with about 10 % of global photosynthetic production (Hohmann-Marriott & Blankenship 2011). Consequently, they play a vital role in the carbon cycle. Furthermore, cyanobacteria play a crucial role in the planetary nitrogen assimilation cycle due to the capability of some of them to fix atmospheric N_2 gas (Karl et al. 1997, Zehr et al. 2001).

2. Cyanobacterial metabolism

Cyanobacteria are photoautotrophic prokaryotes that can perform oxygenic photosynthesis and fix atmospheric carbon dioxide via Calvin-Benson cycle for primary metabolism (Hohmann-Marriott & Blankenship 2011, Burnap et al. 2015). The nitrogen assimilation via the GS/GOGAT (glutamine synthase/glutamine oxoglutarate aminotransferase) cycle leads to a consumption of 2-oxoglutarate (2-OG), which represents an indicator of the intracellular carbon/nitrogen balance (Muro-Pastor et al. 2001 & 2005, Forchhammer 2004, Commichau et al. 2006, Luque & Forchhammer 2008, Huergo & Dixon 2015). The carbon/nitrogen assimilation reactions via the Calvin-Benson and GS/GOGAT cycles require a tight regulation and a constant sensing of the quantity and quality of the carbon/nitrogen availability. Generally, the nitrogen and carbon metabolisms are coordinated by a complex crosstalk between different input signals (Fig. 1) (Commichau et al. 2006, Luque & Forchhammer 2008). The sensing and regulation of the nitrogen/carbon metabolism in cyanobacteria mainly depends on the signal-transduction protein PII, which senses the energy/carbon/nitrogen status of the cell through binding ATP/ADP and 2-OG in presence of ATP (Fig. 1) (Forchhammer 2008, Forchhammer 2010, Chellamuthu et al. 2013, Huergo et al. 2013, Espinosa et al. 2014, Forchhammer & Lüddecke 2016).

3. Cyanobacterial carbon metabolism

In cyanobacteria, the carbon acquisition takes place mainly via CO_2 fixation reactions. The RubisCO (ribulose-1,5-bisphosphate carboxylase/oxygenase) is the central CO_2 fixation enzyme and uses both, CO_2 (carboxylase activity) and O_2 (oxygenase activity), as substrates. The RubisCO carboxylation reaction generates two molecules of 3-phosphoglycerate (3-PGA), which is further used in the Calvin-Benson cycle, while the RubisCO oxygenase activity releases one molecule of 3-PGA and one molecule of 2-phosphoglycolate (2-PG) (Hohmann-Marriott & Blankenship 2011, Burnap et al. 2015). The 2-PG is toxic to cyanobacteria and has a huge negative effect on the RubisCO and the Calvin-Benson cycle. The detoxification of 2-PG can be

achieved by recycling of 2-PG into 3-PGA only via energetic photorespiration reactions (Eisenhut et al. 2008, Hagemann et al. 2010). In cyanobacteria, the low substrate specificity of RubisCO to discriminate between CO_2/O_2 with the low availability of C_i sources, renders the oxygenase activity highly probable. To overcome the low affinity of RubisCO, cyanobacteria evolved the CCM to elevate the C_i levels around RubisCO carboxylating sites (Burnap et al. 2015).

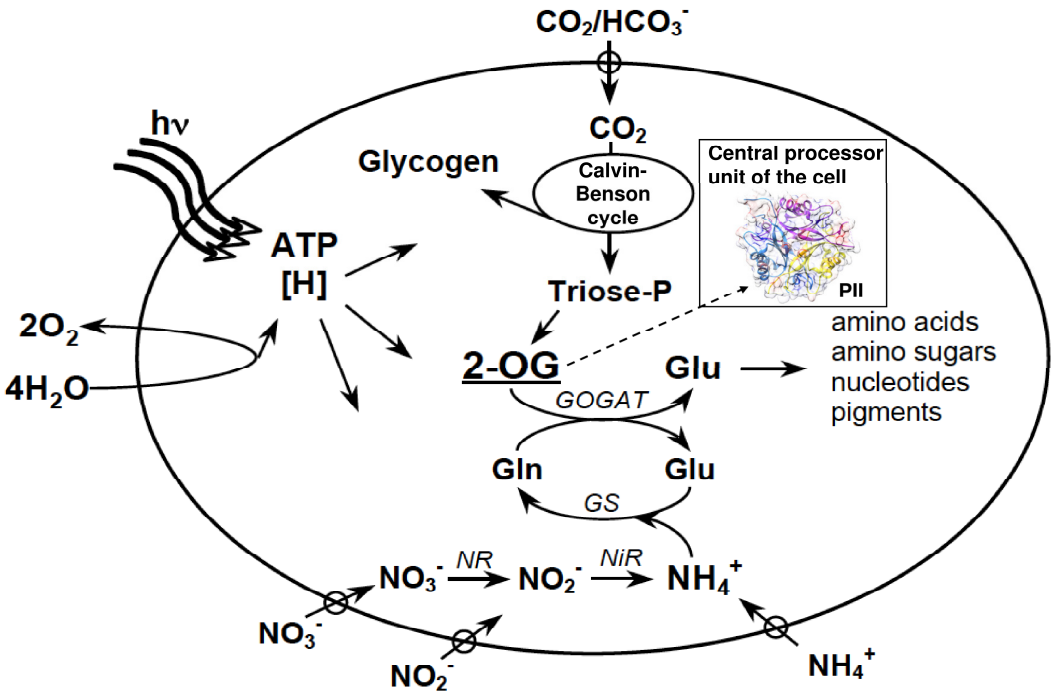


Fig. 1. The crosstalk between nitrogen and carbon assimilation reactions in non-diazotrophic cyanobacteria is modulated by cell signaling PII protein, modified from (Luque & Forchhammer 2008).

The cyanobacterial CCM is expressed at a basal level under CO_2-rich conditions, but under C_i limitation, the CCM expression and activity are maximal. However, a constitutive induction of CCM is also observed even under C_i excess conditions (Wang et al. 2004, Rae et al. 2013, Burnap et al. 2015). Cyanobacterial CCM is composed of several uptake systems for active CO_2 and HCO_3^- transport, as the environmental C_i exists mainly as dynamic interconvertible forms between HCO_3^- and CO_2, and a special organelle called carboxysome, where RubisCO and CA enzymes are present.

Overall, in cyanobacteria five CCM-uptake systems have been identified so far, among them three HCO_3^- transporters and two CO_2 uptake systems (Fig. 2). The cyanobacterial HCO_3^- uptake systems consist of: a high affinity ATP-dependent ABC-type transporter (BCT1 complex) which includes CmpA/B/C/D subunits encoded by the *cmpABCD* operon (Omata et al. 1999), and two

sodium dependent bicarbonate transporters: 1) a low C_i inducible/high affinity transporter SbtA (Shibata et al. 2002) and 2) a constitutive medium affinity/high flux transporter BicA (Fig. 2) (Price et al. 2004, Burnap et al. 2013, Long et al. 2016). Moreover, the cyanobacterial CO_2 uptake systems are organized into two thylakoid-bound systems: 1) the high affinity/low C_i-inducible redox driven CO_2 uptake NDH-I_3 complex encoded by ndhF3/ndhD3/chpY (cupA) subunits; and 2) the low affinity/constitutive redox driven CO_2 uptake NDH-I_4 complex encoded by ndhF4/ndhD4/chpX (cupB) subunits (Fig. 2) (Orf et al. 2015). The CO_2-NDH-$I_{3/4}$ complexes convert CO_2 into HCO_3^- by hydrating CO_2 and creating a proton gradient. The NDH-$I_{3/4}$ systems are NADPH- (or Ferredoxin) dependent to rise the proton gradient, thereby increasing the ATP levels for CO_2 assimilation and for promoting further hydration of CO_2 (Ma & Ogawa 2015). By this mechanism, NDH-$I_{3/4}$ are able to convert CO_2 into HCO_3^-, albeit the reaction equilibrium lies rather on the side of CO_2, since the intracellular HCO_3^- concentration in the cytosol is higher. Hence, the CO_2-NDH-$I_{3/4}$ systems keep a low CO_2 diffusion gradient by trapping the diffused CO_2 in form of HCO_3^-, so that further CO_2 diffusion into the cell is possible (Shibata et al. 2001, Daley et al. 2012, Long et al. 2016).

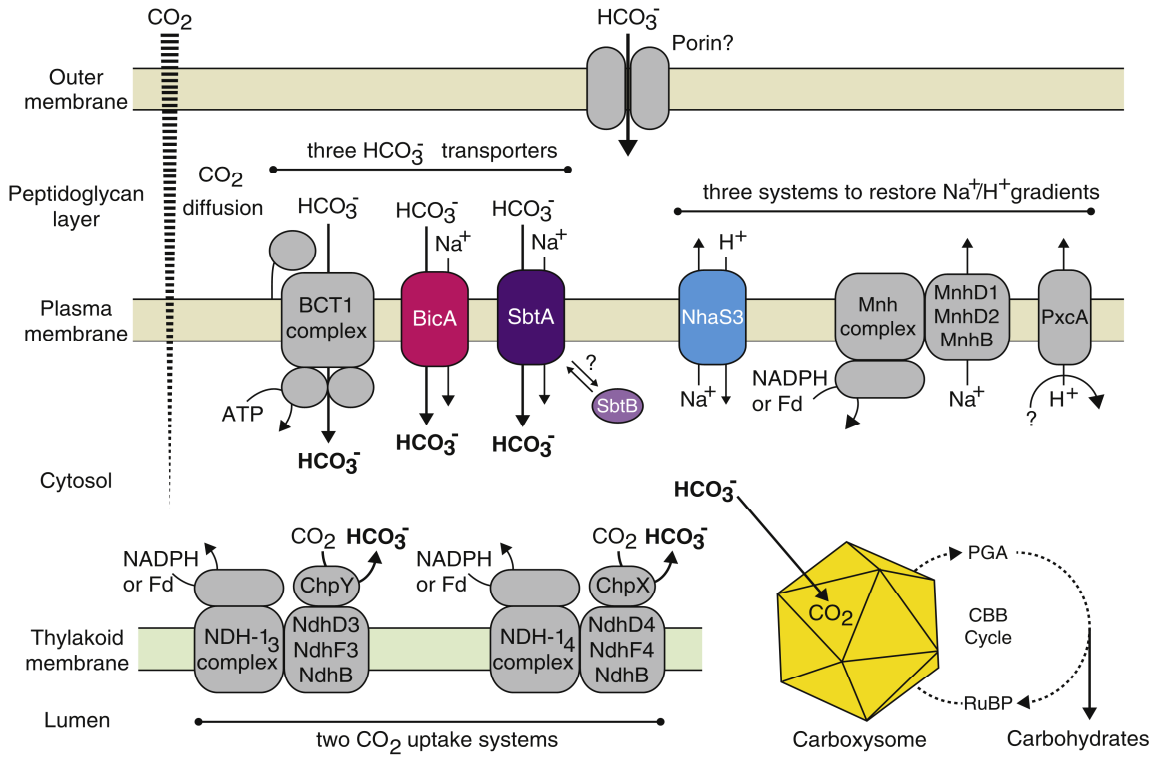


Fig. 2. Schematic representation of cyanobacterial CCM, adapted from (Long et al. 2016).

Hence, the active transport of the entire CCM permits to condensate C_i in cyanobacterial cytosol to 1000-fold higher levels than in the extracellular space (Daley et al. 2012). The accumulated HCO_3^- is then utilized within the carboxysome, where HCO_3^- is dehydrated to CO_2 by CA and then CO_2 fixed into organic carbon by RubisCO (Long et al. 2016). The carboxysomes are cyanobacterial microcompartments with characteristic positively charged icosahedral structure, which probably attracts the negatively charged HCO_3^- (Long et al. 2007, Tsai et al. 2007). Notably, in the cyanobacterium *Synechocystis* sp., the number of carboxysomes is also enhanced under low C_i conditions (Hackenberg et al. 2012). The localization of CA and RubisCO together within the carboxysomes ensures the rise of the local concentrations of C_i in the vicinity of RubisCO and thereby minimizing the RubisCO-oxygenation reactions. Physiological studies revealed that carboxysomes are required for a healthy growth of cyanobacteria and for full function of RubisCO (Price & Badger 1991, Burnap et al. 2015, Long et al. 2016).

Interestingly, the HCO_3^- transporter *sbtA* is located in a bicistronic operon next to a second gene, encoding for a putative regulatory protein of SbtA transporter, SbtB (Srivastava et al. 2005, Du et al. 2014, Burnap et al. 2015, Long et al. 2016). The *sbtB* gene product (a PII-like paralogue) belongs to the widespread cell signaling proteins of the PII superfamily (Forchhammer & Lüddecke 2016). Transcriptional analysis of *Synechocystis* sp. PCC 6803 and *Synechococcus elongatus* PCC 7942 revealed that SbtB is expressed with SbtA under C_i limitation, and the *Sbt* operon is completely repressed under enriched CO_2 conditions (Wang et al. 2004, Schwarz et al. 2011). Moreover, a phosphoproteome study of *Synechocystis* sp. PCC 6803 implied that SbtA is a phosphoprotein, whose phosphorylation level is down regulated, while SbtB protein is highly phosphorylated on Thr53 (i.e. resembling posttranslational phosphorylation of PII protein in cyanobacteria), under nitrogen starvation conditions (Spät et al. 2015, Forchhammer & Lüddecke 2016). In *E. coli*, the heterologous expression of SbtB protein along with SbtA revealed that SbtB was able to bind to SbtA and inhibited the HCO_3^- uptake activity of SbtA, implying that SbtB is likely a regulator of SbtA (Du et al. 2014).

Three transcriptional factors, namely NdhR (CcmR), CmpR and CyAbrB2 have been identified to be responsible for the transcriptional regulation of C_i uptake system and other genes (Figge et al. 2001, Wang et al. 2004, Takahashi et al. 2004, Klähn et al. 2015, Orf et al. 2016). The transcription factors NdhR and CmpR belong to the LysR family and are regulated by small molecular mass metabolites such as 2-PG or 2-oxoglutarate (2-OG). The 2-PG and ribulose 1,5-bisphosphate are co-activators for CmpR and NdhR (Nishimura et al. 2008, Jiang et al. 2018), while 2-OG or $NADP^+$ are serving as co-repressors for NdhR (Daley et al. 2012). In addition to these metabolic signals, the accumulated cellular bicarbonate pool seems to play a repressing role on low C_i-induced

genes of cyanobacteria (Woodger et al. 2005, Orf et al. 2015). Remarkably, it was shown that soluble adenylyl cyclase (sAC) activity, including that of the enzymes from the cyanobacteria *Synechocystis* sp. and *Anabaena* sp. PCC 7120 (Hammer et al. 2006), is regulated by varying bicarbonate concentrations (Chen et al. 2000, Steegborn et al. 2005**)**. The stimulation of sAC activity through sensing the cellular HCO_3^- level seems to be universal and evolutionary conserved across different phyla separated by millions of years (Chen et al. 2000). These findings led to the hypothesis that sAC could play a role as evolutionarily conserved HCO_3^- sensor in prokaryotes and eukaryotes to detect the fluctuation in C_i and accordingly modulates the level of the second messenger cyclic-AMP (cAMP) (Chen et al. 2000, Raven 2006). The cAMP is considered as one of the most prevalent second messengers in nature, playing fundamental roles in regulating many crucial physiological processes, which also includes the well-understood response of enterobacteria to different organic carbon sources (Cann et al. 2003, Steegborn et al. 2005, Gancedo 2013). Among cyanobacteria, cAMP has been shown to be involved in respiration, light sensing and cell motility (Agostoni & Montgomery 2014).

4. Signal transduction PII proteins

4.1. Discovery and classification of PII proteins

The PII signal transduction family is widely spread in all domains of life, where they represent one of the largest and most ancient family of signaling proteins in nature. The PII protein has been discovered around 50 years ago in the end of 60s of the last century. The discovery of the PII protein is one of the classical biochemical stories. The initial work on the GS revealed that the activity of GS can be modulated by unknown protein(s). Using gel filtration approach, two fractions of were isolated and designated PI and PII. From the fraction PI, the adenylyltransferase (ATase) enzyme, catalyzing the adenylylation/de-adenylylation of GS, was purified to high homogeneity. From fraction PII, one protein was required to regulate the ATase activity, and since then and up to now the PII protein has kept its simple term of the PII fraction (Shapiro 1969). The PII proteins play a crucial role in the regulation of anabolic reactions of nitrogen as well as carbon metabolism of many bacteria, algae and land plants (Sant'Anna et al. 2009; Huergo et al. 2013, Chellamuthu et al. 2013). The PII protein exerts its functions through regulation of many enzymes, transport channels, and transcription factors (Ninfa & Jiang 2005, Commichau et al. 2006, Forchhammer 2008, Sant'Anna et al. 2009, Forchhammer 2010, Chellamuthu et al. 2013, Huergo et al. 2013, Forchhammer & Lüddecke 2016).

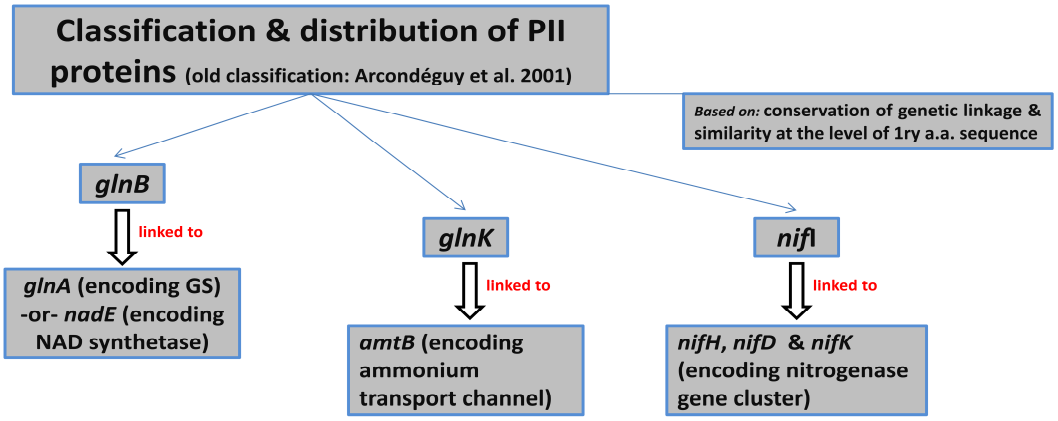
The primary classification of the PII proteins family has been achieved by Arcondéguy et al. (2001) based on the conservation of genetic linkage to the subfamilies of PII proteins and the similarity at the primary amino acid sequences (Fig. 3). The proteins of PII family classified into three sub-groups: 1) PII of *glnB*-subfamily (linked to GS or NAD synthetase), 2) PII of *glnK*-subfamily (linked to ammonium transport channel AmtB), and 3) PII of *nifI*-subfamily (linked to nitrogenase enzymes cluster). With release of more genomic data, Sant'Anna and his collages recently revised the PII family and grouped PII of *glnB*/*K* together in one subfamily (Fig. 3). The new classification of PII family revealed an extra subfamily (called PII-new group; PII-NG) of so far uncharacterized PII proteins lacking PROSITE signature (PS00496 & PS00638; encoding for uridylylation sites & T-loop of PII proteins) in genetic association with heavy metals efflux pumps (Fig. 3), therefore, they proposed a role in heavy metal sensing (Sant'Anna et al. 2009). The evolutionary tracing of the PII family suggested that PII proteins of the *glnK*-subfamily is the ancestor of the PII family, as it seems that they evolved early with amtB in the prokaryotes, most likely in Archaea, and later by horizontal gene transfer events relocated into the deep branching lineages of prokaryotic domains. Then subsequently, by gene duplication events, PII proteins of the *glnB*-subfamily would have been arisen. Finally, through an endosymbiosis events, the eukaryotic lineages of Plantae would have been gained the PII proteins later from cyanobacteria (Fig. 3) (Sant'Anna et al. 2009).

4.2. Sensing properties of canonical PII proteins

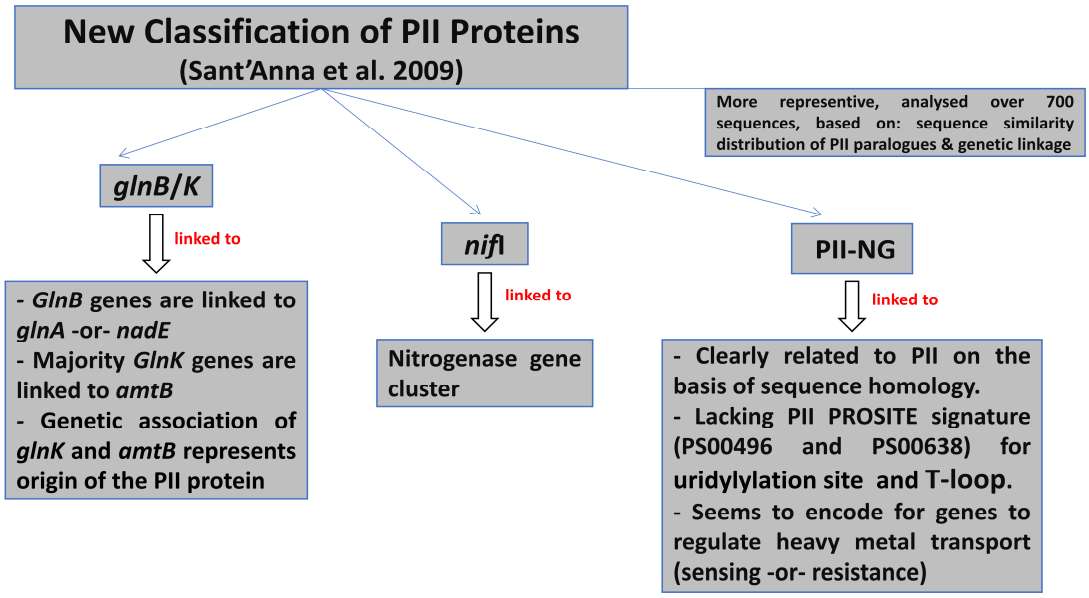
PII proteins have a fundamental role as energy/carbon/nitrogen sensors. The binding of ATP/ADP and 2-OG to PII proteins permit the cell to estimate the current energy and nitrogen/carbon status. Accordingly, the binding of the small effector molecules, allosterically to PII proteins, causes signaling events through conformational changes within the PII body, which in turn enables the PII proteins to bind to different targets (transcription factors, key metabolic enzymes and transporters) to regulate the current metabolic situation (Fokina et al. 2010a&b, Truan et al. 2010, Radchenko & Merrick 2011, Zeth et al. 2014, Forchhammer & Lüddecke 2016).

Overall, PII proteins sense the cellular energy state by binding of ATP or ADP competitively to the same binding site, and sense C/N state of the cell by binding 2-OG (Smith et al. 2003, Fokina et al. 2010a, Oliveira et al. 2015, Forchhammer & Lüddecke 2016). 2-OG, an intermediate of the TCA-cycle, is used as the carbon skeleton for nitrogen incorporation by the GS/GOGAT-cycle. Therefore, the level of 2-OG indicates the current nitrogen/carbon balance of the cell (Smith et al. 2003, Fokina et al. 2010a, Oliveira et al. 2015).

A



B



C

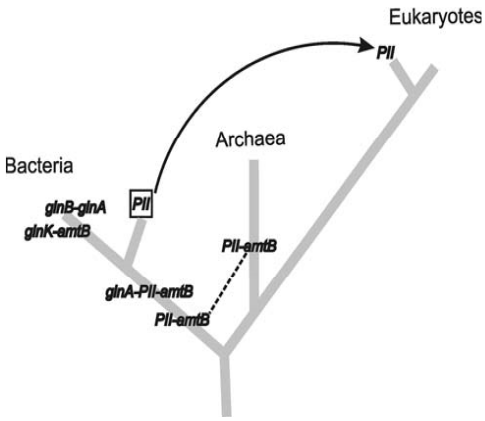


Fig. 3. Classification and evolution of cell signaling proteins of PII family. (A) old classification according to Arcondéguy et al. (2001). (B) new classification according to Sant'Anna et al. (2009). (C) Evolution of PII family across the tree of life. PII (GlnK) was originated in Proteobacteria and horizontally transferred across Prokaryote (represented by dash line). Later *glnB* originated by gene duplication events. Cyanobacterial PII (represented by a square) transferred to Eukaryotes (Plantae) via endosymbiosis (represented by a black arrow), (C) is adapted from (Sant'Anna et al. 2009).

Under conditions of high 2-OG levels (poor nitrogen supply), the ATP-dependent PII binding to 2-OG causes strong conformational changes in a surface exposed element of PII called T-loop (see below), which in turn impair the interaction of PII proteins with different targets like in case of N-acetyl-L-glutamate kinase (NAGK) (Forchhammer 2008, 2010) and ammonium transport channel AmtB (Radchenko et al. 2010). Hence, the binding of 2-OG adds extra secondary regulation for PII proteins (Forchhammer 2008).

Representatives of PII proteins are present in almost all bacteria (including cyanobacteria), in nitrogen-fixing archaea (Arcondéguy et al. 2001; Forchhammer 2008) and in eukaryotic oxygenic phototrophs (Uhrig et al. 2009). In eukaryotes, PII homologues have only been identified in Chloroplastida (land plants and green algae), where they are nuclear-encoded, and in Rhodophyta, where they are plastid-encoded (Uhrig et al. 2009). In both these groups, PII is localized in the chloroplast (Hsieh et al. 1998, Ermilova et al. 2013). In Chloroplastida, PII proteins control the ornithine pathway via activity regulation of NAGK, as in cyanobacteria (Chellamuthu et al. 2014). It is noteworthy that in green algae and land plants NAGK activity responds to the cellular glutamine levels via PII signaling (Chellamuthu et al. 2014, Minaeva et al. 2015). In the evolution of the Chloroplastida, PII acquired a new signaling function through the evolution of a C-terminal extension, called Q-loop, which forms a glutamine binding site (Chellamuthu et al. 2014). Interestingly, the PII protein from *Arabidopsis thaliana*, is exceptional because it is able to bind 2-OG in presence of ATP and ADP with less affinity in presence of ADP (Smith et al. 2003), and it does not respond to glutamine like other plant PII proteins, due to a modification of the Q-loop (Chellamuthu et al. 2014). Therefore, it seems that the PII proteins in the *Brassicaceae* family, to which *A. thaliana* belongs, evolved in a different way than in other plants.

In cyanobacterium *S. elongatus* PCC 7942, PII protein is encoded by *glnB* gene and binds to the known PII effector molecules (ATP, ADP, and 2-OG) (Fokina et al. 2010b). The regulation of PII proteins through binding small effector molecules (ATP, ADP, and 2-OG) has been described as well for many other bacteria including *Escherichia coli*, *Azospirillum brasilense*, *Herbaspirillum seropedicae* and *Mycobacterium tuberculosis*, for different archaeal species *Haloferax mediterranei* and *Archaeoglobus fulgidus*, and for Chloroplastida *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* (Smith et al. 2003, Jiang & Ninfa 2007, Bandyopadhyay et al. 2010, Fokina et al. 2010a&b, Helfmann et al. 2010, Truan et al. 2010, Maier et al. 2011, Chellamuthu et et al. 2014, Palanca et al. 2014, Oliveira et al. 2015).

In general, PII proteins possess 3 binding sites between the inter-cleft of three subunits of PII trimer. The anti-cooperativity binding of the effector molecules was reported for cyanobacterial PII protein from *S. elongatus* PCC 7942, to sense different cellular concentrations of the cellular

metabolites (ATP/ADP and 2-OG) (Fokina et al. 2010a&b, Ma et al. 2013). The saturation of the first binding site causes subtle conformational changes in the subsequent binding sites, leading to meaningful decrease in the affinities of the other binding sites with increasing the K_d values (Fokina et al. 2010a&b, Ma et al. 2013, Zeth et al. 2014). Notably, the affinity of PII proteins to bind ATP is higher than ADP, and moreover, 2-OG antagonizes the ADP binding by increasing the binding affinity of the competing nucleotide ATP (Fokina et al. 2010b, 2011, Radchenko et al. 2010, da Rocha et al. 2013). Hence, the 2-OG and ATP binding to PII proteins are cooperative, while ADP does not support the 2-OG binding (Smith et al. 2003, Fokina et al. 2010b, Gerhardt et al. 2012, Oliveira et al. 2015). A recent study proposed an ATPase switch between ATP to ADP bound statue with a weak ATPase activity of PII proteins, which is inhibited by 2-OG (Radchenko et al. 2013). However, researchers ruled out later the ability of cyanobacterial PII protein to switch between the interacting targets with ATPase activity of PII (Lüddecke & Forchhammer 2015). Additionally, a simulation model suggested that PII with fully ADP-bound status is unlikely to exist under standard physiological conditions (da Rocha et al. 2013). Finally, a new study raised the possibility of AMP as an important signaling molecule in archaea, where the authors identified the ability of PII protein from an archaeon *Haloferax mediterranei* to bind ATP, ADP and AMP (Palanca et al. 2014).

4.3. Posttranslational modifications of PII proteins

Remarkably, the signaling properties of canonical PII proteins depend not only on the response to the fluctuation of the effector molecules, but also on the covalent modification of S49/Y51 resides of the T-loop. In proteobacteria, including *E. coli*, the Y51 is subject to uridylylation under nitrogen limiting conditions (Merrick 2015, Forchhammer & Lüddecke 2016). The uridylylation reaction is catalysed by a bifunctional enzyme, encoded by *GlnD*, possessing uridylyltransferase (for PII uridylylation) and uridylyl-removase (for PII deuridylylation) activities (Atkinson et al. 1994, Jiang et al. 1998, Zhang et al. 2010, Merrick 2015, Forchhammer & Lüddecke 2016). The GlnD binds to Gln and its activity is controlled by the intracellular glutamine concentrations. Under nitrogen rich conditions, in response to the increase of Gln levels, the Gln binding to GlnD enhances GlnD-uridylyl-removase activity. While, under nitrogen starvation, the drop in Gln levels enhances the GlnD-uridylyltransferase to uridylylate PII-Y51 (Merrick 2015, Forchhammer & Lüddecke 2016). Notably, GlnD enzyme is present ubiquitously in the actinobacteria and the proteobacteria, however in actinobacteria, the GlnD activity adenylylates PII at Y51, rather than uridylylating it (Hesketh et al. 2002, Strösser et al. 2004, Huergo et al. 2013, Merrick 2015, Forchhammer & Lüddecke 2016). Therefore, the posttranslational modification of Y51 governs

PII interactions such that, the GlnD enzymes behave as Gln gauge using PII as signal transmitter (for more details, see below PII-targets) (Jiang et al. 2012, Forchhammer & Lüddecke 2016).
In unicellular cyanobacteria *Synechocystis* sp. PCC 6803 and *S. elongatus* PCC 7942, the T-loop of PII proteins is phosphorylated at S49 residue in a sequential manner in response to nitrogen limitation (Forchhammer & Tandeau de Marsac 1994, Forchhammer 2004, Forchhammer & Lüddecke 2016). PII phosphorylation is probably induced by elevated levels of 2-OG under nitrogen-deprived conditions, however the molecular mechanism underlaying the PII phosphorylation remains enigmatic as the PII-kinase is still unknown so far (Merrick 2015, Forchhammer & Lüddecke 2016). Interestingly, a recent work suggested that PII phosphorylation is not achieved by a specific PII-kinase but by multiple kinases, where they can replace each other equally (Spät 2017). However, the dephosphorylation of PII proteins under nitrogen rich conditions is driven by a protein of PP2C phosphatase family, named PphA (Irmler & Forchhammer 2001, Ruppert et al. 2002, Kloft et al. 2005, Su et al. 2011, Su & Forchhammer 2013). The PphA activity is repressed when PII is bound to ATP/2-OG (Irmler et al. 1997, Kloft et al. 2005, Su et al. 2011, Su & Forchhammer 2013, Merrick 2015, Forchhammer & Lüddecke 2016). Thus, 2-OG plays a significant role as a signaling molecule, both by promoting PII signaling via direct binding to PII under nitrogen starvation conditions and also by repressing the PphA activity to assure that PII stays phosphorylated (Merrick 2015). Interestingly, in some cyanobacteria, like in *Nostoc* and *Prochlorococcus*, the T-loop phosphorylation site S49 is not conserved (Forchhammer et al. 2004, Merrick 2015). In the filamentous cyanobacterium *Nostoc* sp. PCC 7120, a novel nitration modification at Y51 has been described (Zhang et al. 2007), even though the physiological relevance of such modification remains elusive so far. Up to now, there is no evidence for covalent modifications of PII proteins in archaea and plants, nor in the Firmicutes such as *Bacillus subtilis* (Smith et al. 2004, Merrick 2015). Recently, experimental data indicated that algal PII protein is nonphosphorylated in *C. reinhardtii* under nitrogen starvation (Ermilova et al. 2013), supporting the idea that plant PII proteins differ markedly than cyanobacterial PII proteins regarding the covalent phosphorylation modification, despite the conservation of phosphorylation site S49 (Chellamuthu et al. 2013, Ermilova et al. 2013).

4.4. PII superfamily (structure and PII-like proteins)

Canonical PII proteins are highly conserved homo-trimeric proteins with three characteristic loop regions (B-, C- and T-loops) (PDB: 2XUL) (Fokina et al. 2010a&b, Truan et al. 2010). Each subunit of PII trimer consists of two βαβ motifs connected between β2 and β3 strands by the long flexible surface exposed T-loop. Therefore, the overall secondary structure arrangement of each

monomer subunit of canonical PII protein is β1-α1-β2-(T-loop)-β3-α2-(B-loop)-β4-(C-loop). Recently, a C-terminal extension in plant PII proteins namely Q-loop has been identified as a glutamine sensor site (PDB: 4USJ) (Chellamuthu et al. 2014). These loop regions are located near the inter-subunit clefts and play a major role in protein-protein interactions and ligand binding properties. PII-like proteins are similar and clearly related to canonical PII proteins but lacking the PII PROSITE signature sequences and the functions for most of them are still unknown (Sant'Anna et al. 2009, Forchhammer & Lüddecke 2016). Bioinformatics and structural genomics approaches expanded the members of the PII superfamily and proposed that the PII-like proteins represent an even more widespread family of trimeric regulators, occurring in almost all living organisms (Sant'Anna et al. 2009, Forchhammer & Lüddecke 2016).

In general, the signaling proteins of the PII superfamily are characterized by their highly conserved trimeric structure, consisting of a triangular core of β-sheets from the ferredoxin-like fold of the three subunits (Fig. 4) (Forchhammer & Lüddecke 2016, Wheatley et al. 2016). Despite the highly conserved structure, the amino acid sequence conservation is quite low, implying that the signaling proteins of PII superfamily are involved in regulation of different cellular activities, which differ markedly from canonical PII proteins (Forchhammer & Lüddecke 2016, Wheatley et al. 2016). Recently, a new member of the PII superfamily has been characterized, the carboxysome-associated PII protein (Fig. 4) (CPII; PDB: 5DS7). CPII protein was found to bind ADP/AMP and bicarbonate and was proposed to sense the bicarbonate availability to control the carbon metabolism in proteobacteria *Thiomonas intermedia* (Wheatley et al. 2016). Furthermore, in firmicutes such as *Staphylococcus aureus* or *Bacillus subtillus*, a PII-like protein (Fig. 4) (termed PstA or DarA; PDB: 4WK1) was identified to sense the secondary messenger c-di-AMP; however, the exact function of that protein is still unknown (Campeotto et al. 2015, Gundlach et al. 2015, Müller et al. 2015).

Another PII-like protein, bearing the architectural core principle of PII proteins with the lowest sequence identity to canonical PII proteins among all PII-like proteins (Fig. 4) (Arnesano et al. 2003, Forchhammer & Lüddecke 2016), is the divalent ion tolerance proteins CutA (PDB: 1NAQ), which are universally distributed in all domains of the life, including bacteria, archaea, plants, protozoa, animals, and humans. The CutA protein was first discovered in *E. coli* in a gene locus involved in divalent metal tolerance. The genetical mutation in the CutA locus in *E. coli* showed increased sensitivity to divalent metal ions (copper, zinc, nickel, cobalt and cadmium) owing to increased heavy metals uptake. Therefore, the CutA gene product was speculated to confer heavy meatal tolerance (Fong et al. 1995, Arnesano et al. 2003). In mammals, CutA protein is localized in the brain cells for processing, trafficking and anchoring the membrane-bound

neurotransmitter enzyme acetylcholinesterase (AChE), besides it is supposed to control the copper homeostasis and ligand transport to membranes (Perrier et al. 2000, Arnesano et al. 2003, Yang et al. 2008, Liang et al. 2009). However, a physiological study revealed that CutA is not essential for Cu^{2+} tolerance/adaptation in knockout lines of *A. thaliana* (Burkhead et al. 2003).

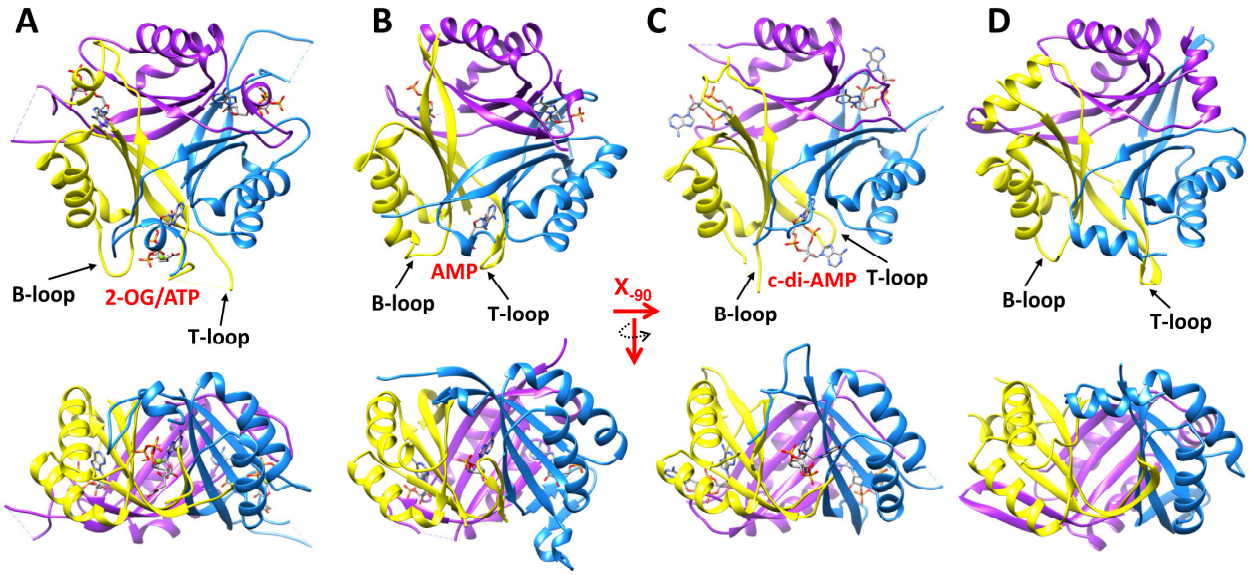


Fig. 4. Structural comparison between different members of PII superfamily revealed conservation of trimeric architecture. Top and side views of (A) canonical PII protein bound to ATP and 2-OG (PDB:2XUL), (B) PII-like protein CPII bound to AMP (PDB: 5DS7), (C) PII-like protein PstA bound to c-di-AMP (PDB: 4WK1), and (D) PII-like protein CutA (PDB: 1NAQ). Structural elements (B- and T-loops) are indicated.

Furthermore, the trimeric PII-like domain was found to be embedded as well within enzymes as a sensor domain to regulate the catalytic domains of the protein (Chellamuthu et al. 2013). The ATP-phosphoribosyltransferase (HisG; PDB: 2VD3 & 1NH8) is the first enzyme in biosynthetic pathway of histidine with C-terminal PII domain. HisG enzymes are allosterically inhibited by histidine and AMP. In the intersubunit clefts of PII-like domain, similar to canonical PII proteins, HisG can bind the allosteric inhibitor to mediate the inhibition of phosphoribosyltransferase activity (Cho et al. 2003, Lohkamp et al. 2004, Chellamuthu et al. 2013). Another example of enzyme bears PII-like domain is NGG1p-interacting factor 3 (NIF3; PDB: 2GX8). In NIF3, the PII-like domain is embedded in the middle of the NIF3 proteins, away from the N- and C-terminal domains (Saikatendu et al. 2006, Godsey et al. 2007, Fujishiro et al. 2014). The function of NIF3 was proposed as an iron chaperone in *Methanocaldococcus jannaschii*, that is involved in the biosynthesis of the iron-guanylylpyridinol cofactor (Fujishiro et al. 2014), however the exact cellular functions of NIF3 are still mysterious.

4.5. The wide range of the PII signaling network (PII interacting targets)

4.5.1. PII targets in oxygenic phototrophs

In oxygenic phototrophs (including: cyanobacteria, algae and land plants), the major interaction partners of PII described so far are the enzyme N-acetyl-L-glutamate kinase (NAGK), the biotin carboxyl carrier protein- (BCCP) subunit of acetyl-CoA carboxylase (ACCase) and restricted to cyanobacteria only; the transcriptional co-activator PipX (Forchhammer & Lüddecke 2016, Forcada-Nadal et al. 2018). NAGK is the controlling enzyme of arginine biosynthesis and PipX is a transcriptional co-activator of the master transcription factor of nitrogen regulated genes, NtcA (Heinrich et al. 2004, Espinosa et al. 2006, Beez et al. 2009, Fokina et al. 2010b, Fokina et al. 2011, Forchhammer & Lüddecke 2016, Forcada-Nadal et al. 2018). PII binds to BCCP-subunit of ACCase, to regulate the fatty acids biosynthesis (Feria-Bourrellier et al. 2010, Gerhardt et al. 2015, Hauf et al. 2016b). Furthermore, physiological studies indicate possible regulation of nitrate uptake as well as ammonium transport (Lee et al. 2000, Kloft & Forchhammer 2005). Regulation of the ammonium transport channel AmtB is well known from various bacteria and archaea (Radchenko et al. 2010), where in most cases, the corresponding PII protein is of *glnK*-subfamily. The network of PII interactions in cyanobacteria is summarized in (Fig. 5).

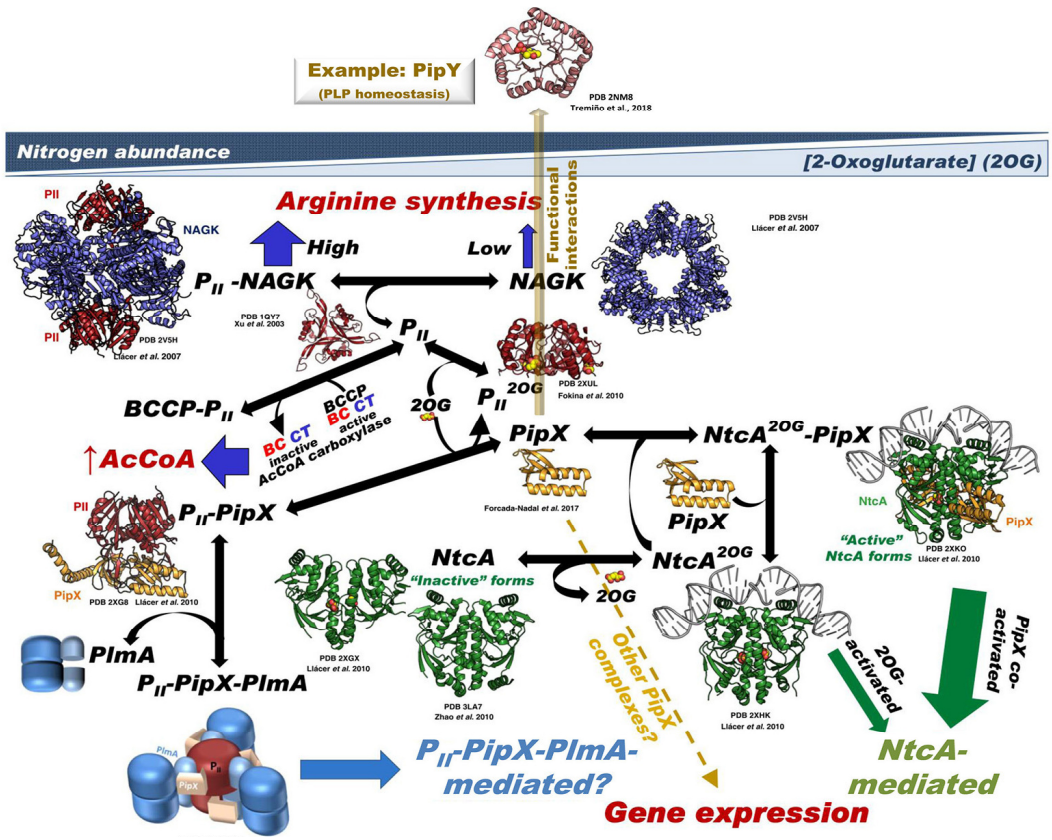


Fig. 5. Network of PII interactions in cyanobacteria, adapted from (Forcada-Nadal et al. 2018).

4.5.1.1. N-acetyl-L-glutamate kinase (NAGK)

NAGK catalysis the phosphorylation of N-acetyl-L-glutamate (NAG) to N-acetyl-L-glutamyl-phosphate (NAG-P), which represents the rate limiting step in the arginine biosynthesis pathway (Caldovic & Tuchman 2003). In oxygenic phototrophs, the NAGK activity is regulated by PII and the feedback inhibitor, arginine (Forchhammer & Lüddecke 2016, Forcada-Nadal et al. 2018).

In case of sufficient energy and nitrogen supply, indicated by a high ATP and a low 2-OG levels, PII forms an ATP-dependent activating complex with NAGK (Fig. 5) (Heinrich et al. 2004, Llácer et al. 2007, Mizuno et al. 2007, Beez et al. 2009, Chellamuthu et al. 2014). The NAGK seems to be a specific target for PII proteins in oxygenic phototrophs only, where the PII-NAGK interaction is extremely conserved during the evolution of oxygenic phototrophs (Beez et al. 2009, Chellamuthu et al. 2013, Forcada-Nadal et al. 2018). Furthermore, several crystal structures of PII-NAGK complexes were reported from: cyanobacterium *S. elongatus* (PDB: 2V5H), plant *Arabidopsis thaliana* (PDB: 2RD5), and a hetero complex for algal *Chlamydomonas reinhardtii* PII over *A. thaliana* NAGK (PDB: 4USJ). Structurally, NAGK forms a hexameric ring by a trimer of dimers, therefore, two PII trimers sandwich the hexameric NAGK (Fig. 6).

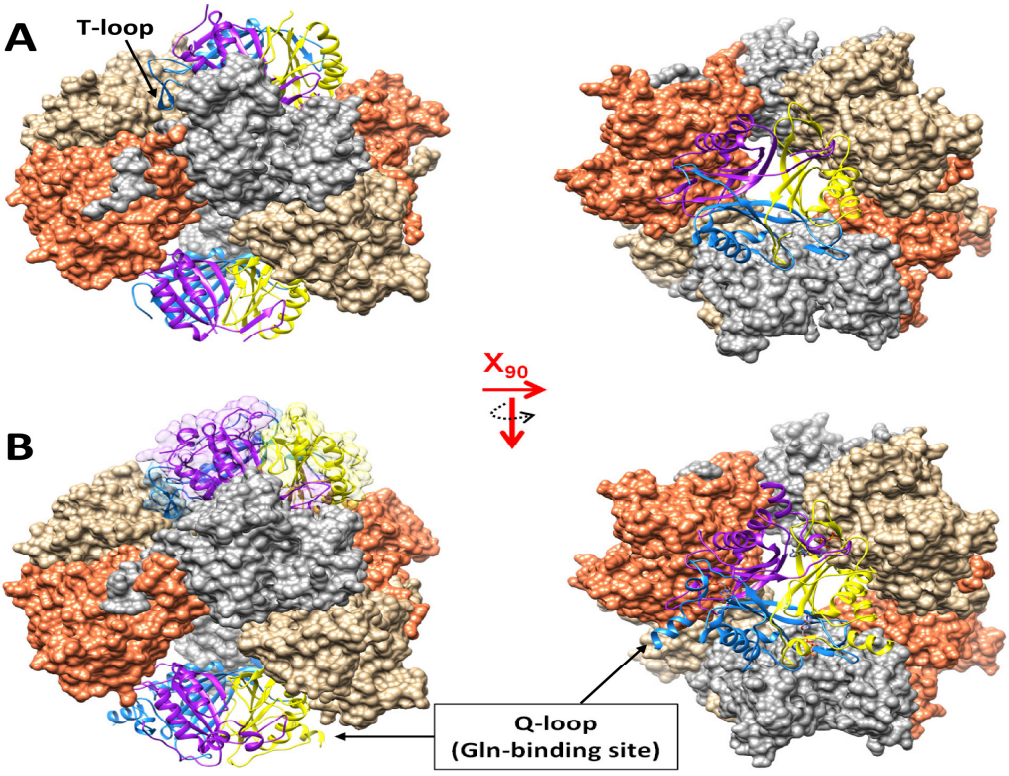


Fig. 6. Structure of PII-NAGK complex. The hexameric NAGK ring (surface representation) is sandwiched by two PII trimers (ribbon representation) and the interaction is mediated by the T-loop of the PII protein. Side and top view of (A) cyanobacterial PII-NAGK complex (PDB: 2V5H), and (B) algal/plant PII-NAGK complex (PDB: 4USJ) showing evolution of the Gln binding site (Q-loop).

Each subunit of the PII trimer interacts with each NAGK subunit via the B- and T-loops, where the T-loop reaches deeply into the interdomain cavity of NAGK (Llácer et al. 2007, Mizuno et al. 2007, Chellamuthu et al. 2014). When NAGK is complexed with PII, the arginine binding sites within NAGK are broadened, causing a decrease in the affinity of NAGK for arginine with an apparent increase of arginine K_d value. Whereas, the PII-NAGK complex formation enhances the affinities for NAG with an apparent decrease of K_m value for NAG (Llácer et al. 2007, Beez et al. 2009, Forcada-Nadal et al. 2018). Consequently, the PII-NAGK complex formation increases the catalytically efficiency of NAGK and decrease the feedback inhibitory effect of arginine to the NAGK-reaction (Heinrich et al. 2004, Beez et al. 2009). Under conditions of high 2-OG levels (poor nitrogen supply), the ATP-dependent binding of 2-OG to PII causes strong conformational changes in the T-loop, which in turn impair the interaction of PII protein with NAGK, leading to complex dissociation and restoring the susceptibility of NAGK to arginine feedback inhibition. Additionally, the N-limitation dependent phosphorylation of PII T-loop at Ser49 prohibits the PII-NAGK interactions (Forchhammer & Lüddecke 2016).

A mechanistic explanation of the PII-NAGK complex formation proposed a two-step mechanism for the complex formation: first, an encounter complex, that involves the B-loop. Here it comes to a contact between R233 of NAGK and E85 of PII. This breaks a salt bridge between E85 and R47 in the T-loop of PII, which is now free and can adopt a bended structure. In the second step, the bended T-loop deeply inserts into the NAGK clefts to form the tight complex (Fokina et al. 2010b). In addition, a PII variant with single amino acid replacement, the Ile-86 mutated to Asn (PDB: 2XBP structure), facilitates the formation of the T-loop bended conformation (i.e. similar to the T-loop confirmation in the complex with NAGK) via a H-bond to T43. Consequently, the PII (I86N) variant can bind constitutively to NAGK *in vitro* and *in vivo*. The strong over activation of NAGK, caused by the PII (I86N) variant, leads to an increased intracellular arginine level, followed by cyanophycin accumulation in the unicellular cyanobacterium *Synechocystis* sp. PCC 6803 (Fokina et al. 2010b, Watzer et al. 2015).

Remarkably, in higher plants, except in *Brassicae* (e.g. *A. thaliana*), the PII-NAGK interaction is Gln sensitive due to the evolution of a PII C-terminal extension (Q-loop), representing a novel glutamine binding site (Fig. 6) (Chellamuthu et al. 2014), while the cyanobacterial and *A. thaliana* PII interacts with NAGK in glutamine-insensitive manner (Chellamuthu et al. 2014).

4.5.1.2. PII interacting protein X (PipX)

Another major interacting partner in the PII signaling network in cyanobacteria is the PII interacting protein X (PipX). The PipX, a protein uniquely found only in cyanobacteria, can interact with PII

and the global transcription factor NtcA, acting as a co-activator of the latter (Espinosa et al. 2007). The nitrogen starvation condition leads to an induction of the NtcA regulon. *In vivo*, the PipX-NtcA interaction is required for NtcA dependent gene expression under nitrogen limiting conditions (Espinosa et al. 2006, 2007, 2014). The cellular limitation of the nitrogen supply causes an increase in the intracellular levels of 2-OG. Under elevated levels of 2-OG, the PII-PipX complex formation is impaired due to ATP dependent binding of 2-OG to PII (Llácer et al. 2010). As a consequence, the free PipX can bind and activate the transcription factor NtcA, causing an induction of the NtcA regulon (Llácer et al. 2010). In case of nitrogen excess conditions, indicated by a low intracellular 2-OG concentration, three PipX monomers can bind to one PII trimer (Fokina et al. 2011). The sequestration of PipX by PII leads to a decreased NtcA-PipX complex formation (Llácer et al. 2010, Espinosa et al. 2014), and thereby deactivation of NtcA (Fig. 5). Hence, this 2-OG depending partner switch of PipX provides a mechanistic link between PII signaling and the NtcA depending gene expression (Espinosa et al. 2014). The crystal structure of the PII-PipX complex (PDB: 2XG8) reflected that the T-loop of PII acts as an antenna that binds at the lateral cleft of PipX subunits (Fig. 7) (Llácer et al. 2010, Zhao et al. 2010). Unlike, the bent conformation of the T-loop in NAGK-PII complex, the T-loop in the PipX-PII complex is in an extended conformation, which is promoted by ADP (Llácer et al. 2010). In addition, *in vitro* studies revealed that ATP and 2-OG have a negative influence on PII-PipX complex formation, while ADP further stabilizes and supports the PII-PipX complex formation (Espinosa et al. 2006, Fokina et al. 2011).

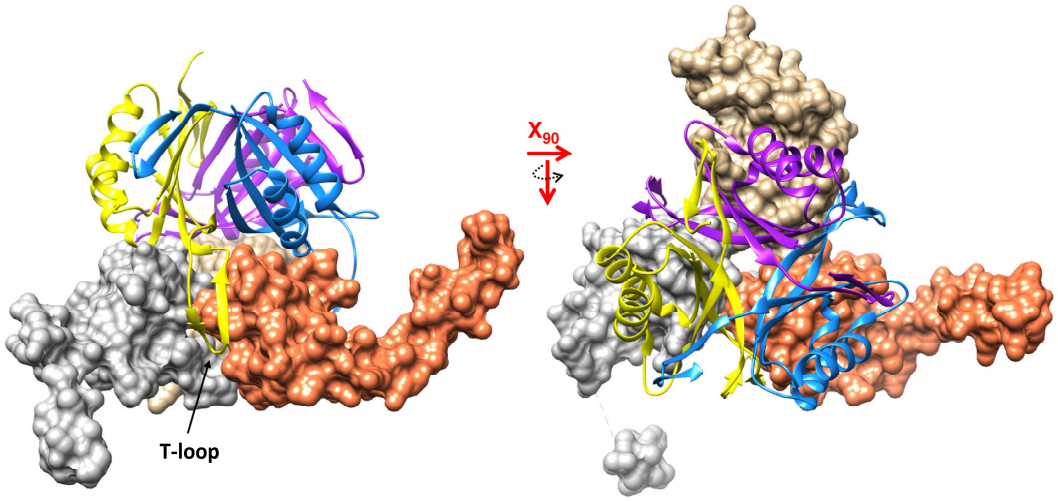


Fig. 7. Structure of the cyanobacterial PII-PipX complex. Side and top view of PII-PipX complex (PDB: 2XG8) reveals that each subunit of the PII trimer (ribbon representation) binds to one PipX (surface representation) and the interaction is mediated by the T-loop of the PII protein.

4.5.1.3. The GntR-like transcription regulator: PlmA

In cyanobacteria, a GntR-like transcription regulator PlmA was identified in a yeast three-hybrid screening to interact concurrently with PII-PipX complex (Fig. 5), while PlmA was not able to interact with PipX or PII alone (Labella et al. 2016). Apparently, PII is unable to interact directly with PlmA, while PII is needed indirectly to induce the extended conformation of PipX found in the PII-PipX complex (Fig. 7) (PDB: 2XG8), which then might facilitate the interaction between PlmA and the C-terminal helix of PipX (Llácer et al. 2010, Labella et al. 2016, Forcada-Nadal et al. 2018). However, the physiological function of the ternary PII-PipX-PlmA remains unclear. Notably, the in-solution NMR structure for PipX revealed that the signaling C-terminal helix of the noncomplexed PipX adopts a flexed confirmation, supporting the notion of importance of PII to induce the PipX extended conformation needed to interact with PlmA (Forcada-Nadal et al. 2017). The attempts to inactivate PlmA in the cyanobacterium *S. elongatus* PCC7942 was not successful, implying that PlmA is essential for *S. elongatus* lifestyle (Labella et al. 2016). However, in other cyanobacteria like *Synechocystis* sp. PCC 6803 and *Anabaena* sp. PCC 7120 and in marine picocyanobacteria, a pleiotropic function of PlmA was proposed to be involved in: 1) photosystem stoichiometry, 2) chlorophyll fluorescent kinetics, 3) plasmid maintenance 4) heterocyst development, and 5) regulation of sRNA YFR2 (Labella et al. 2016, Forcada-Nadal et al. 2018).

4.5.1.4. Biotin carboxyl carrier protein- (BCCP) subunit of acetyl-CoA carboxylase (ACCase)

The ACCase enzyme controls the rate limiting step of fatty acid biosynthesis. The bacterial ACCase is composed of three functional subunits: 1) the biotin carboxylase (BC; homodimer), 2) the biotin carboxyl carrier protein (BCCP); and 3) the carboxyltransferase (CT, heterotetramer) (Gerhardt et al. 2015). The plant and bacterial (including cyanobacteria) PII (GlnB) proteins are able to complex with the BCCP-subunit of ACCase to tune down the enzyme's activity (Feria-Bourrellier et al. 2010, Rodrigues et al. 2014, Gerhardt et al. 2015, Hauf et al. 2016b). Accordingly, the PII-BCCP complex formation inhibits the fatty acid biosynthesis by decreasing the ACCase turnover, and thereby acetyl-CoA can be directed to other pathways (Fig. 5) (Gerhardt et al. 2015, Forcada-Nadal et al. 2018). *In vitro*, PII-BCCP interaction is maximized in the presence of ATP, while 2-OG prohibits the complex formation and promotes the complex dissociation (Gerhardt et al. 2015, Hauf et al. 2016b). Mutational analysis revealed that the PII T-loop could mediate the PII-BCCP interaction (Hauf et al. 2016b). *In vivo* studies indicated that the fatty acids biosynthesis

and the lipid bodies formation are enhanced in a PII deletion mutant (Δ*PII*) of the cyanobacterium *Synechocystis* sp. PCC6803 and *E. coli*, due to the lack of PII control on the ACCase activity (Hauf et al. 2016b, Rodrigues et al. 2018). The model for the role of PII in modulation of the ACCase activity is summarized in (Fig. 8).

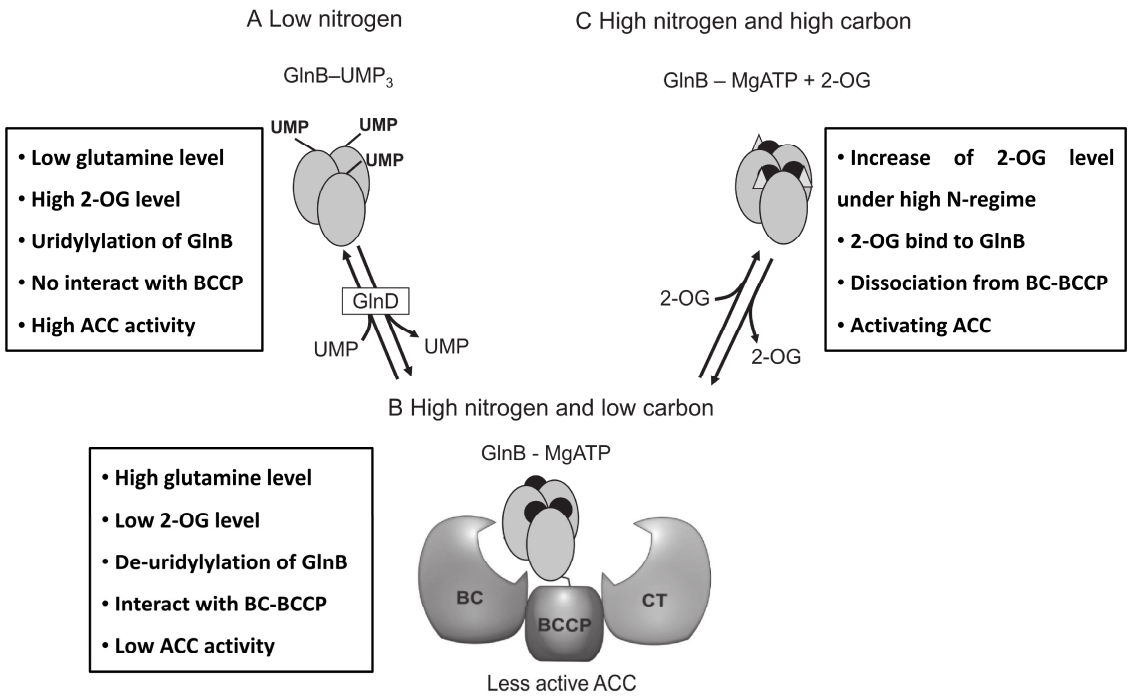


Fig. 8. Model for the mechanistic role of the PII protein in the regulation of ACC activity through binding to the BCCP subunit of the ACCase enzyme, modified from (Gerhardt et al. 2015).

4.5.1.5. PII associated membrane protein A (PamA)

In the cyanobacterium *Synechocystis* PCC 6803, the PII associated membrane protein A (PamA), encoded by the open reading frame *sll0985* was identified by yeast two-hybrid assays as a target of PII interaction. The PamA is a member of transmembrane mechanosensitive channels of the MscS subfamily (mechanosensitive channel of small conductance) (Forchhammer 2010). The *in vitro* pulldown analysis proved the PII-PamA interaction in an ATP/2-OG-sensitive manner (Osanai et al. 2005). Therefore, it was speculated that PII regulates the transport activity of PamA to close/open the channel under nitrogen-replete/depletion conditions, respectively, according to the availability of 2-OG (i.e. the ATP-dependent binding of 2-OG to PII causes the dissociation of the PII-PamA complex to open the channel). However, the physiological substrate transported by PamA remains ambiguous so far. Physiological studies revealed that PamA is implicated in sugar

and nitrogen metabolism in *Synechocystis* PCC 6803, whereas the mutation of PamA causes a decrease in the transcript of sugar catabolic genes and several nitrogen NtcA-regulated genes (Osanai et al. 2005, Forchhammer 2008). Besides, the Δ*pamA* mutant was not able to grow in glucose. Notably, the PII-PamA interaction does not seem to be preserved among cyanobacteria, because the PII protein from *S. elongatus* failed to interact with the putative *S. elongatus* PamA protein and moreover the PamA protein is not conserved in cyanobacteria (Forcada-Nadal et al. 2018).

In term of abundancy of all known proteins to interact with the PII protein in the cyanobacterium *S. elongatus*, PII is by far the most abundant protein and after the approximate sum of all identified PII-network within cyanobacteria, still about 80% of PII protein would be free, which raises the possibility of additional unidentified targets of the PII-network (Fig. 9). Indeed, proteomic-based pulldown analysis of PII protein from cyanobacteria *S. elongatus* and *Synechocystis* sp. PCC6803 revealed that PII proteins could interact with several new targets, but at present the relevance of such finding waits for further biochemical and physiological confirmations (Spät 2017).

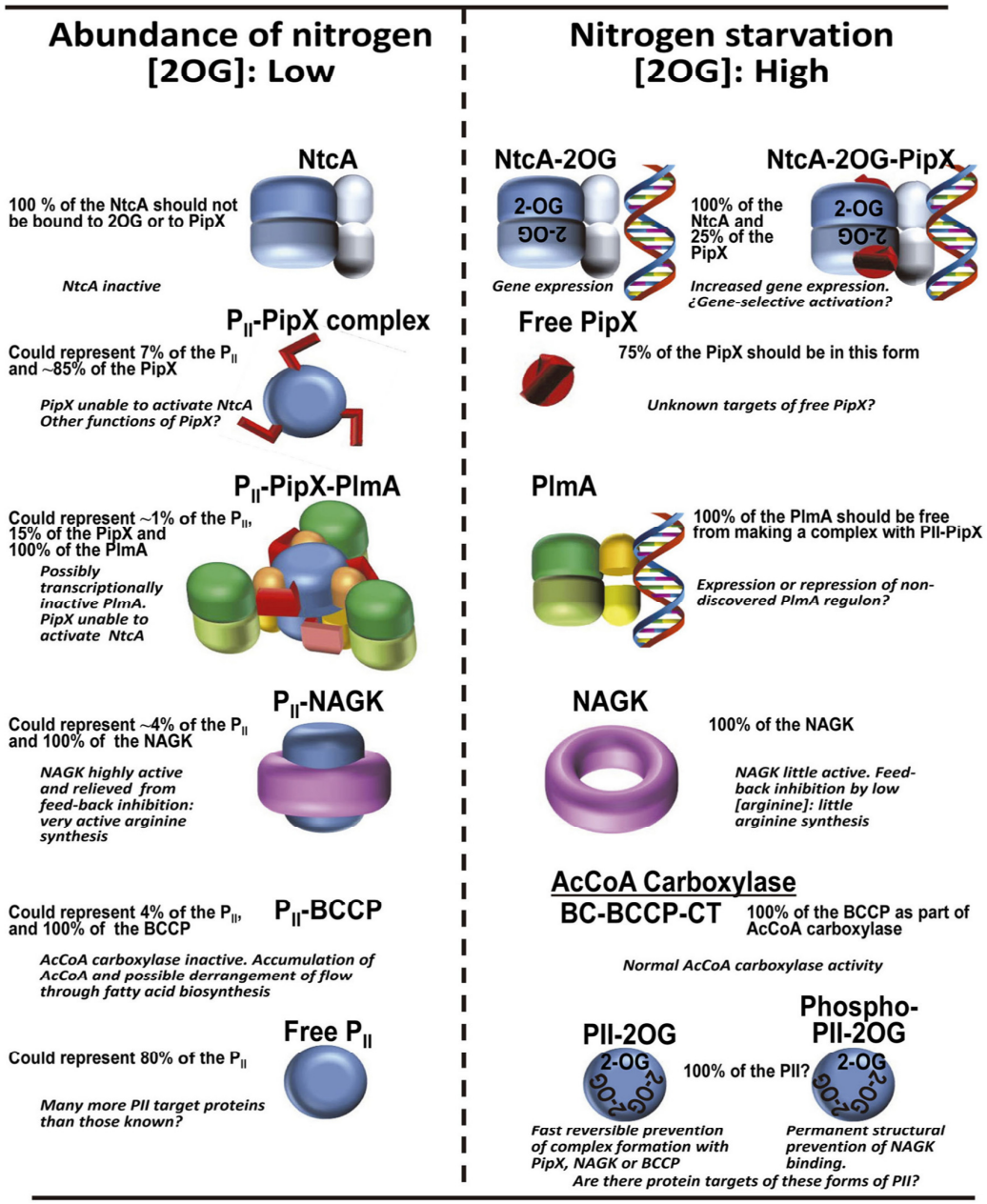


Fig. 9. Relative estimated amount of bound/free PII protein within identified PII-network in *S. elongatus*, adapted from (Forcada-Nadal et al. 2018).

4.5.2. Others bacterial PII targets

4.5.2.1. Ammonium transport channel AmtB

One of the most important nitrogen sources for archaea and bacteria is ammonium (NH_4^+). In archaea and bacteria, NH_4^+ is transported via the homotrimeric ammonium transport channel AmtB (PDB: 1U7G), which is consistently linked to PII of the glnK-subfamily (Coutts et al. 2002, Khademi et al. 2004). To prevent excess accumulation of intracellular NH_4^+, PII (GlnK) negatively

regulates the AmtB channel via direct interaction with AmtB, thereby closing the channel and inhibiting the AmtB transport activity (Coutts et al. 2002, Conroy et al. 2007, Radchenko et al. 2010). The crystal structure of the inhibitory AmtB-PII complex from *E. coli* revealed that PII-AmtB interaction is mediated mainly via the surface exposed PII T-loop, where the tips of the T-loops insert into the exit pores of the trimeric-AmtB to block the channel (PDB: 2NUU) (Fig. 10) (Conroy et al. 2007). Under nitrogen excess condition of NH_4^+ shock, NH_4^+ is taken up actively via the AmtB channel, which can be used to aminate Glu by glutamine synthetase (GS), leading to an increase in the intracellular level of Gln and a drop in the intracellular levels of 2-OG and the ATP/ADP ratio. The dramatic changes in the intracellular pool of Gln and 2-OG promotes the de-uridylylation of PII (GlnK) (Radchenko et al. 2010, Radchenko et al. 2014). Consequently, the fully de-uridylylated PII (GlnK) binds ADP, which supports AmtB-PII complex formation to close the AmtB channel (Conroy et al. 2007, Radchenko et al. 2010, Radchenko et al. 2014).

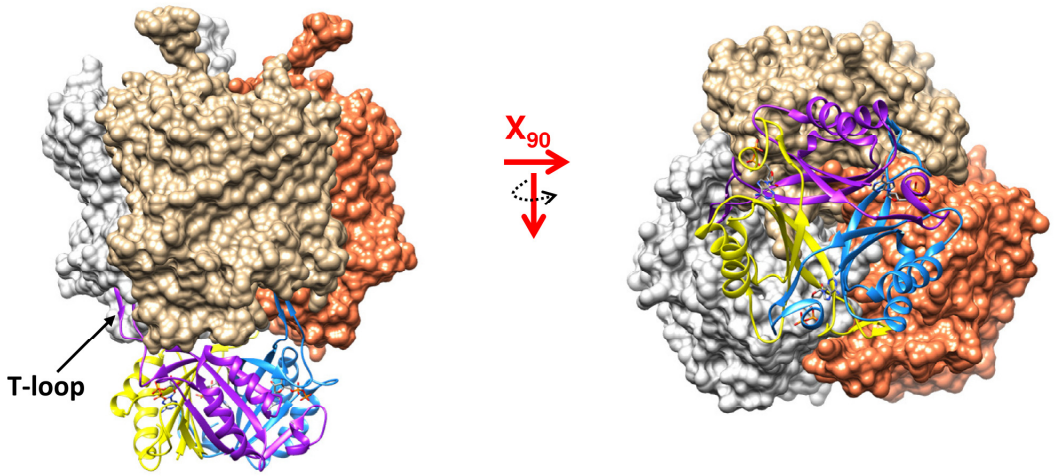


Fig. 10. Structure of PII-AmtB complex. Side and top view of the PII trimer (ribbon representation) shows the closing of the trimeric AmtB channel (surface representation) and that the interaction is mediated by the T-loop of the PII protein (PDB: 2NUU).

This process is entirely reversible, by dropping the extracellular concentrations of NH_4^+ (under nitrogen limiting conditions), the intracellular concentrations of Gln decrease while 2-OG levels and the ATP/ADP ratio rise, leading to a uridylylation of PII (GlnK). Under this condition, the ATP-dependent binding of 2-OG can compete/replace the ADP-bound on PII (GlnK), causing conformational changes in the T-loop and thereby the dissociation of the AmtB-PII complex, thus opening the AmtB channel again. Additionally, the posttranslational uridylylation of PII Y51 prohibits the re-association of PII with AmtB (Radchenko et al. 2010, Radchenko et al. 2014). The

mechanism of PII-(GlnK) dependent activation/inactivation of ammonium transport channel AmtB is summarized in (Fig. 11) (Radchenko et al. 2010).

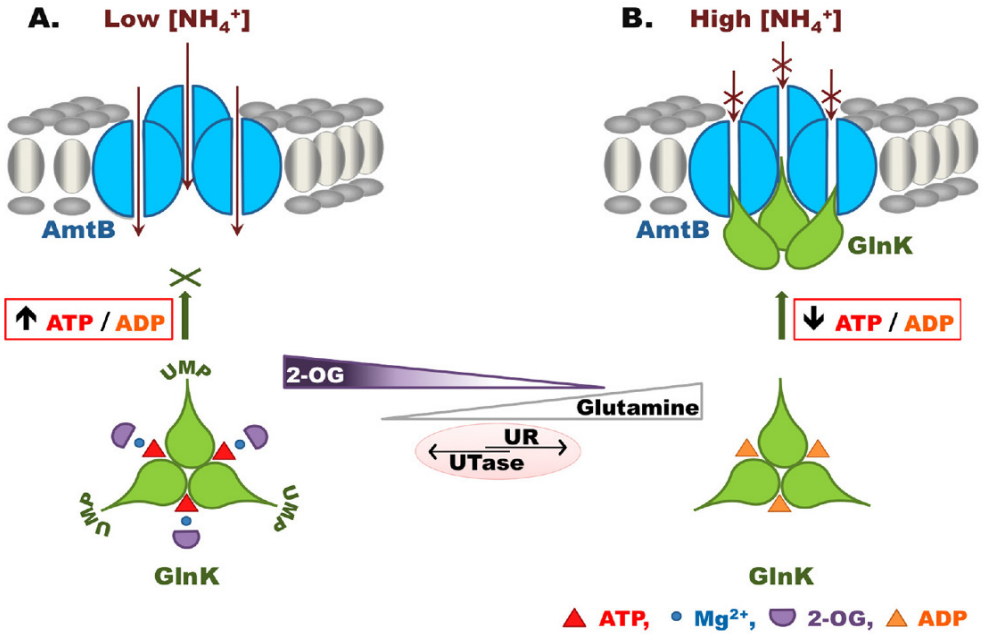


Fig. 11. Model for the mechanistic role of the PII (GlnK) protein in blocking the NH_4^+ transport channel AmtB, adapted from (Radchenko et al. 2010). Under nitrogen limiting conditions (low NH_4^+), the uridylylation of PII is achieved by the high uridylyltransferase (UTase) activity of GlnD enzyme. While, under nitrogen excess conditions (high NH_4^+), the de-uridylylation of PII is accomplished by the uridylyl-removing (UR) activity of GlnD.

In the nitrogen fixing bacterium *Azospirillum brasilense*, the interaction of PII proteins (GlnB and GlnZ) with AmtB is promoted in the presence of adenylated nucleotides (ATP, ADP, and AMP), while the ATP-dependent 2-OG binding encourages the dissociation of the PII-AmtB complex (Huergo et al. 2007). Furthermore, physiological studies indicate the possible regulation of nitrate/nitrite uptake (Lee et al. 2000, Kloft & Forchhammer 2005) as well as ammonium transport in cyanobacteria and plants (Zhang et al. 2007, Ferrario-Méry et al. 2008, Forchhammer 2010, Ohashi et al. 2011).

Interestingly, in *Bacillus subtilis*, the PII protein of GlnK subfamily (*Bs*PII) was found to form a tight membrane localized complex with the AmtB transport channel under both, nitrogen rich (NH_4^+ shock) and limiting (NO_3^- or Glu) conditions (Detsch & Stülke 2003, Heinrich et al. 2006). The membrane dissociation of *Bs*PII protein was induced by ATP and enhanced in the presence of 2-OG; notably *Bs*PII was not able to bind ADP (Heinrich et al. 2006). However, the physiological studies revealed that *Bs*PII prevents the intracellular leakage of the NH_4^+ as well from the *B. subtilis* cells under nitrogen limiting conditions (Fedorova et al. 2013). Whether the leakage

certainly happens through the AmtB channel and the direct binding of *Bs*PII to the AmtB exit pores closes the channels in order to stop the leakage remains enigmatic so far (Fedorova et al. 2013).

4.5.2.2. Nitrogenase regulatory enzymes (DraT and DraG)

In the diazotrophic bacterium *A. brasilense*, PII proteins play a key role in the post-translational regulation of nitrogenase activity. Other main interacting partners of the PII signaling system are the DraT (di-nitrogenase reductase ADP-ribosyl-transferase) and DraG (di-nitrogenase reductase glycol-hydrolase), the nitrogenase regulatory enzymes (Huergo et al. 2006a, 2009). Under nitrogen rich conditions (NH_4^+ shock), the inactivation of the di-nitrogenase reductase (NifH) occurs by ADP ribosylation (ADP-R) with the DraT enzyme. Under nitrogen-fixing conditions (nitrogen limiting conditions), the activation of the NifH occurs by the removal of ADP-R with the DraG enzyme, allowing the nitrogenase activity. Under high levels of extracellular nitrogen, the PII protein of GlnB-gene product interacts and activates the DraT to promote the NifH inactivation, while the PII protein of GlnZ-gene product (a GlnK orthologue) binds to DraG to inactivate it. The PII (GlnZ)-dependent DraG inactivation includes the formation of a ternary membrane localized AmtB-PII-DraG complex in which PII binds concurrently with AmtB and DraG (Huergo et al. 2006b, 2007, 2009, Rajendran et al. 2011). The crystal structure of the PII-DraG complex (PDB: 3O5T) revealed that each monomer at the lateral face of the PII-trimer binds to a monomer of DraG, and the PII interaction-contacts with DraG does not include the PII T-loop (Fig. 12). Then the free T-loop can be used to interact with the AmtB channel, allowing PII to bind simultaneously to AmtB and DraG, forming the ternary AmtB-PII-DraG complex (Fig. 12) (Rajendran et al. 2011).

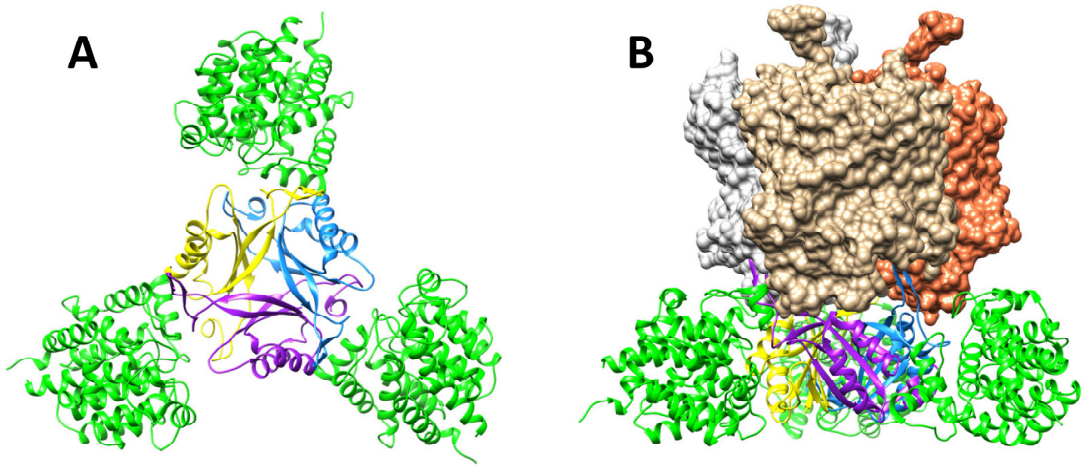


Fig. 12. Structure of PII-DraG, and a model of the ternary AmtB-PII-DraG complex. (A) Top view of PII-DraG complex shows that the interaction of each PII submit (ribbon representation in yellow, violet and blue) with DraG (ribbon representation in green) is not mediated by the T-loop (PDB: 3O5T). (B) Model of the ternary AmtB-PII-DraG complex, compare to Fig. 10 (PII-AmtB complex).

The biochemical analysis revealed that ADP enhances and maximizes the interactions of PII proteins (GlnB and GlnZ) with DraT and DraG, however, the complexes are sensitive to 2-OG, which promotes the dissociation of the complexes. Additionally, the posttranslational modification of PII proteins (uridylylation of Tyr-51) was shown to play a role in the interaction with DraT and DraG, where the binding between PII and DraT/DraG targets is maximized in the presence of de-uridylylated forms of PII (Huergo et al. 2006a & b, 2007, 2009). The mechanism of PII-(GlnB and GlnZ) dependent activation/inactivation of NifH is summarized in (Fig. 13) (Huergo et al. 2009).

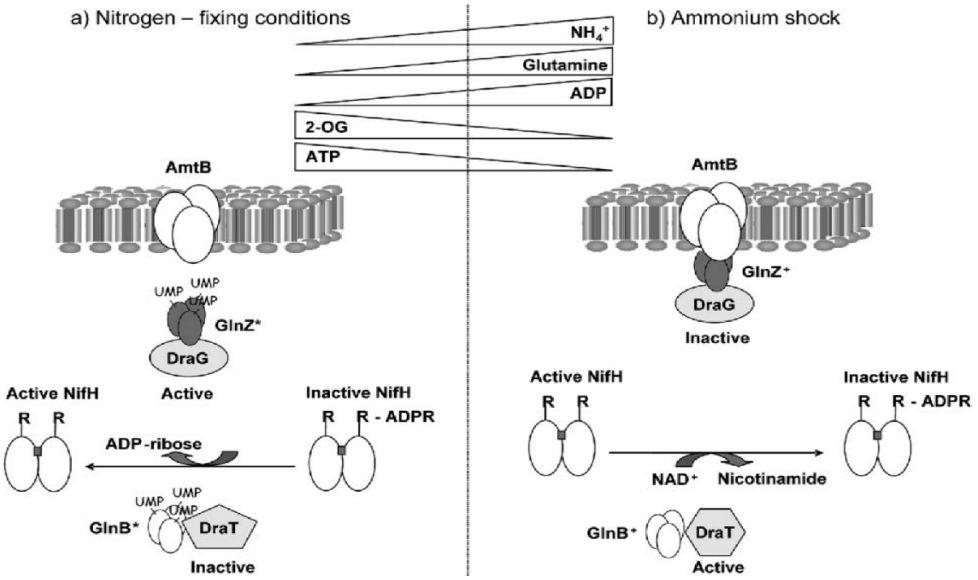


Fig. 13. Model for the mechanistic role of PII (GlnB and GlnZ) proteins in the activation/inactivation of NifH nitrogenase activity, through the interaction with DraT and DraG, adapted from (Huergo et al. 2009). Under nitrogen-fixing conditions, the cytoplasmic localized PII (GlnB and GlnZ) proteins are fully uridylylated and DraG is active, while DraT is inactive, allowing NifH nitrogenase activity. Upon nitrogen excess conditions of NH_4^+ shock, the intracellular increase of Gln due to GS activity leads to de-uridylylation of the PII protein. Under this condition, the de-uridylylated PII (GlnZ) moves to the membrane to close the NH_4^+ transport channel AmtB, and the DraG is sequestered to the membrane through the formation of the ternary AmtB-PII-DraG complex, while DraT starts again to be active leading to ADP ribosylation of NifH (the nitrogenase inactive form). The membrane localization of DraG within the ternary complex, separates the DraG from the cytoplasmic NifH and thereby inhibits the DraG ADP-R removing activity. * indicates fully uridylylated and ATP/2-OG bound PII (GlnB and GlnZ) proteins bound, while + indicates de-uridylylated/ADP bound PII (GlnB and GlnZ) proteins.

4.5.2.3. Glutamine synthetase adenylyl-transferase (ATase)

Another well studied PII interacting partner is the bi-functional enzyme glutamine synthetase adenylyl-transferase (ATase), encoded by the *glnE* gene. The ATase processes an N-terminal

adenylyl-removing (AR) domain and a C-terminal adenylyl-transferase (AT) domain, that catalyses both de-adenylylation/adenylylation of glutamine synthetase (GS), respectively (Huergo et al. 2013). In many bacteria, including the best-studied model organism *E. coli*, the adenylylation of GS leads to inactivation of GS, hence, the ATase imposes posttranslational regulation on the GS activity according to the nitrogen availability. On the other hand, the ATase activity is controlled by both PII proteins (GlnB and GlnK), but PII of GlnB subfamily seems to be more efficient in the regulation of ATase than GlnK *in vitro*. When nitrogen availability is limited, Gln levels drops while 2-OG levels increase. Under this condition, PII (GlnB) is fully uridylylated and saturated with ATP/2-OG. The uridylylated PII (GlnB) protein interacts with the AT domain of the ATase, leading to inhibition of AT activity and promotion of AR activity, thereby enhancing the GS activity (Jiang & Ninfa 2007, 2009, Jiang et al. 2007, Huergo et al. 2013). Under nitrogen rich condition, Gln levels rise dramatically and 2-OG levels drop, while the de-uridylylated form of PII is occupied predominantly by ADP. This form of PII the protein interacts with the AR domain, whereas the Gln binds as well to the AT domain of the ATase enzyme. The binding of PII to the AR domain inhibits the AR activity and stimulates the AT activity, thereby promoting the adenylylation of GS to inactivate it. In addition, the binding of Gln to the AT domain of ATase stimulates the AT activity, stabilizes the de-uridylylated PII-ATase interaction, and prohibits the uridylylated PII form, if any, to complex with the AT domain (Jiang & Ninfa 2007, 2009, Jiang et al. 2007, Huergo et al. 2013). The model of PII-(GlnB) dependent activation/inactivation of GS is summarized in (Fig. 14) (Lüddecke 2016).

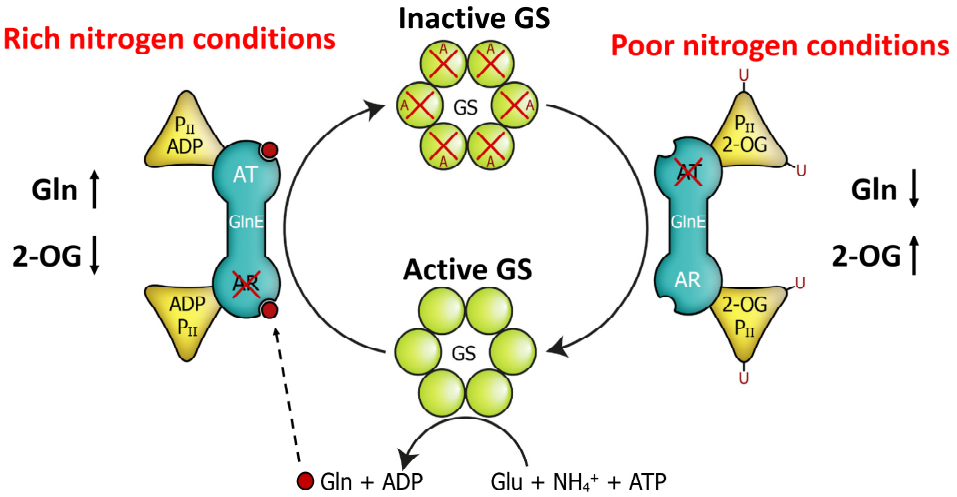


Fig. 14. PII (GlnK) dependent regulation of GS through activation/inactivation of adenylyl-transferase (AT) and adenylyl-removing (AR) activities of the ATase enzyme (GlnE), according to the nitrogen availability in *E. coli*, modified from (Lüddecke 2016).

4.5.2.4. Two-component histidine kinase regulatory system (NtrB/NtrC)

The regulon of the two-component histidine kinase (NtrB/NtrC) system has been well studied in *Proteobacteria*, especially in *E. coli*. The NtrC is the master transcription factor required for the transcription regulation of several nitrogen-related genes, among which the GS and PII (GlnK) proteins are encoded. The NtrC possesses an AAA^+ ATPase domain that facilitates the transcription initiation through the interaction with σ^N RNA polymerase and the C-terminal DNA binding domain. The activity of NtrC is controlled via the phosphorylation/de-phosphorylation by NtrB, a bi-functional enzyme, which possesses an N-terminal phosphatase/phosphotransferase domain and a C-terminal kinase domain. The phosphorylation of NtrC by NtrB phosphotransferase/kinase activity leads to the activation of NtrC regulon via enhancing NtrC-DNA binding affinities and ATPase activity, while the de-phosphorylation of NtrC by NtrB phosphatase activity causes the termination of the NtrC transcription activation (Huergo et al. 2013).

On the other hand, NtrB is regulated by direct protein-protein interaction with PII (GlnB) protein. Under nitrogen-replete conditions, the PII (GlnB) is de-uridylylated, this form of PII can interact with NtrB (Jiang & Ninfa 2007, 2009, Huergo et al. 2013). The PII-NtrB interaction stimulates the NtrB-phosphatase activity and inhibits the NtrB kinase domain and NtrB autophosphorylation activity, leading to a de-phosphorylation of NtrC and accordingly tuning down the gene transcription of the NtrC regulon (Jiang & Ninfa 2007, 2009). In contrast, when the availability of nitrogen sources becomes limited, the 2-OG pools rises, which diminishes progressively the ability of PII to suppress the NtrB autophosphorylation activity. Additionally, under this condition, PII starts gradually to be uridylylated, which facilitates the dissociation of the PII-NtrB complex. Then, the kinase activity of NtrB takes over again, causing a phosphorylation of NtrC and a release of NtrC regulon (Jiang & Ninfa 2009, Huergo et al. 2013). The mechanism of PII-(GlnB) dependent activation/inactivation of the two-component histidine kinase (NtrB/NtrC) system is summarized in (Fig. 15) (Jiang & Ninfa 2009).

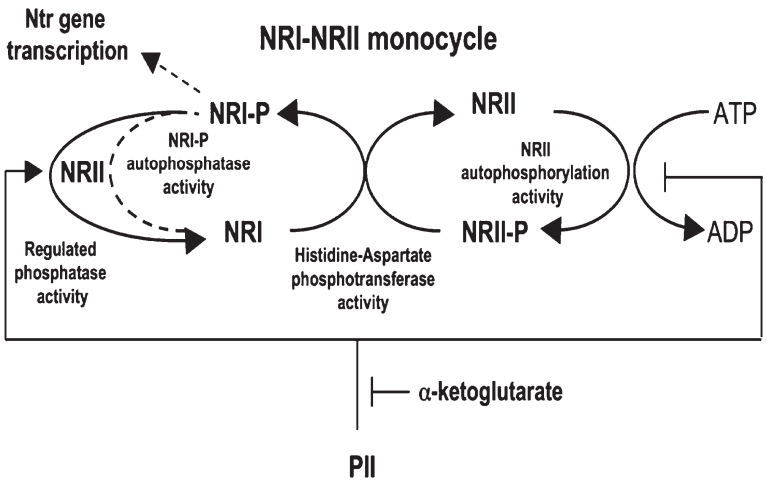


Fig. 15. Model for PII-(GlnB) dependent activation/inactivation of the two-component histidine kinase (NtrB/NtrC) system, adapted from (Jiang & Ninfa 2009).

4.5.2.5. Nitrogen regulatory transcription factors TnrA and GlnR

In *Bacillus subtilis*, the TnrA and GlnR are the master transcription factors of nitrogen assimilation (Fedorova et al. 2013, Schumacher et al. 2015). TnrA can bind *Bs*PII protein and GS enzyme, while the GS can bind simultaneously to the transcription factors TnrA and GlnR. The activity of GS has been shown to be controlled by direct protein–protein interaction with TnrA (Kayumov et al. 2008, Fedorova et al. 2013, Schumacher et al. 2015). Under nitrogen excess conditions, the high intercellular level of Gln promotes the feedback inhibition of GS, then the feedback inhibited GS can interact with TnrA and GlnR (Fedorova et al. 2013, Schumacher et al. 2015). The binding of TnrA to GS improves the effect of Gln-feedback inhibition leading to partial inhibition of GS activity (i.e. the inactive state of GS), while the GS-complexed TnrA is unable to bind to DNA anymore (Fedorova et al. 2013, Schumacher et al. 2015, Hauf et al. 2016a). Whereas, the binding of GS to GlnR causes the activation of the transcription regulon of GlnR (Schumacher et al. 2015). Under nitrogen depleted conditions (NO_3^-), *in vivo* and *in vitro* studies revealed that TnrA is membrane localized and binds to the AmtB-*Bs*PII complex via *Bs*PII in the absence of ATP (Heinrich et al. 2006, Kayumov et al. 2008). Under this condition, the free-unbound TnrA is subject to proteolysis, therefore it seems that the AmtB-*Bs*PII-TnrA complex sequesters TnrA away from the cytoplasmic GS enzyme and protects it from proteolytic hydrolysis (Kayumov et al. 2008, Fedorova et al. 2013). Notably, the AmtB-*Bs*PII-TnrA complex was susceptible as well to ATP and 2-OG binding (Heinrich et al. 2006). Moreover, the cytoplasmic soluble *Bs*PII protein can interact with TnrA as well, forming a stable soluble complex to prevent the interaction of TnrA with the active GS (non-feedback inhibited), causing the release of TnrA regulon (Kayumov et al. 2011,

Fedorova et al. 2013). Recently, the small-angle X-ray scattering analysis (SAXS) of the PII-TnrA complex revealed that each subunit of the PII-trimer can interact with one TrnA dimer (Fig. 16). Strikingly, the interaction between TnrA and PII does not comprise the T-loop but rather the α-helices of PII, explaining the ability of TrnA to form a ternary membrane bound complex as well with AmtB-PII in which PII T-loop is involved in the interaction with the AmtB channel (Conroy et al. 2007, Schumacher et al. 2015). The SAXS model predicted that the ATP binding within the clefts of subunits interferes with the TrnA binding, explaining the ability of ATP to disrupt the PII-TrnA complex (Heinrich et al. 2006, Kayumov et al. 2011, Schumacher et al. 2015). The proposed model of *Bs*PII-(GlnK) dependent activation/inactivation of TnrA is summarized in (Fig. 16) (Schumacher et al. 2015).

Given the similarities between GlnR and TrnA proteins (structurally and DNA-binding mode) (Schumacher et al. 2015), therefore it was speculated that PII would bind also with GlnR (Huergo et al. 2013). Indeed, the PII protein of GlnK subfamily was able to interact with GlnR *in vitro*. The PII-GlnR complex enhanced the DNA binding affinity of GlnR, nevertheless, the physiological importance of such complex remains mysterious (Castellen et al. 2011, Huergo et al. 2013).

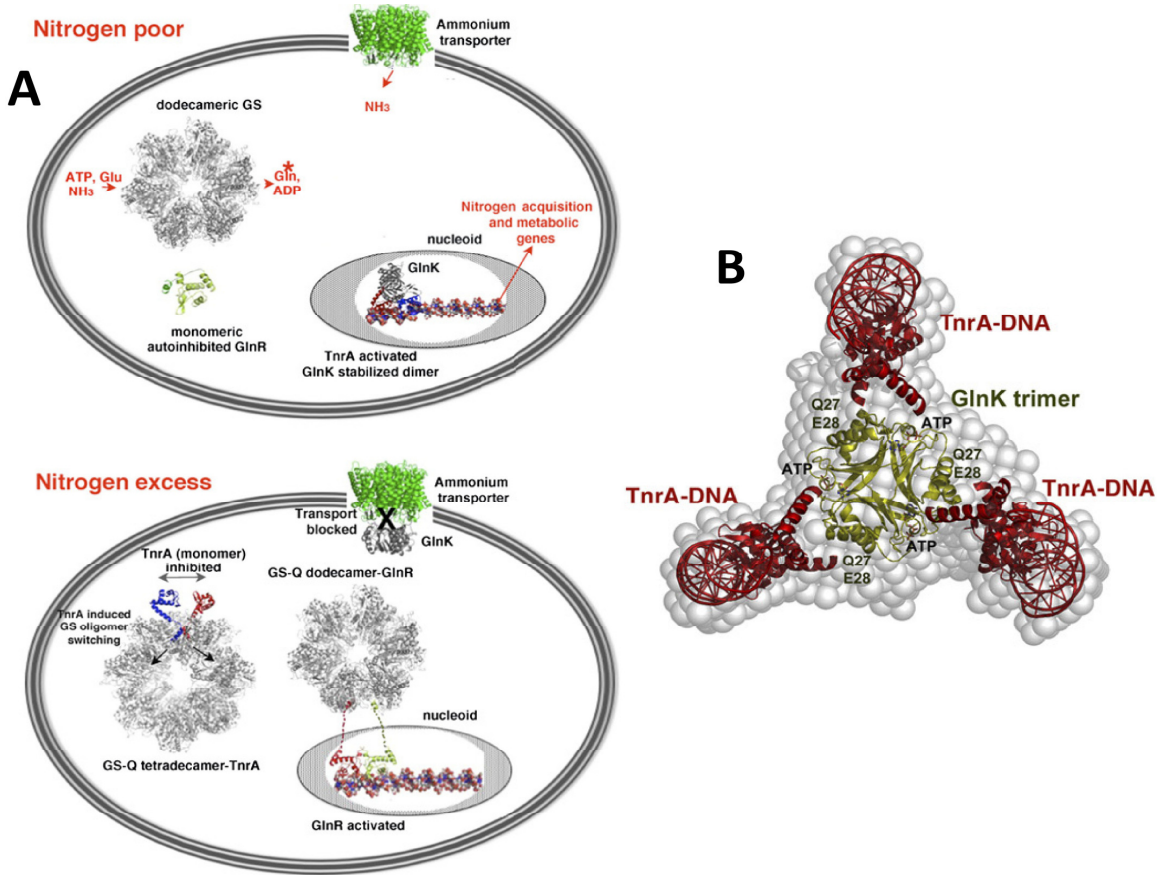


Fig. 16. Model for (A) PII-dependent activation/inactivation of TnrA and (B) the PII-TnrA complex from SAXS analysis, adapted from (Schumacher et al. 2015).

4.5.2.6. Glucosamine 6-phosphate deaminase (NagB)

In *E. coli*, the glucosamine 6-phosphate deaminase (NagB) interactome network revealed the PII protein can interact with and allosterically activate NagB. NagB catalyses the last step in the amino sugar metabolism by cleaving the amide bond of glucosamine 6-phosphate to release ammonia and fructose-6-phosphate. Strictly, only the uridylylated form of PII can interact with NagB and increase its activity to 10-fold in the presence of the positive allosteric regulator, N-acetylglucosamine-6-phosphate (Rodionova et al. 2018). Mechanistically, as discussed above, under nitrogen limiting condition, the PII is fully uridylylated and saturated with high levels of ATP/2-OG causing de-adenylylation of the GS enzyme via the adenylyl-transferase activity of GlnE for maximum activation of GS. Under this condition, the utilization of amino sugars as an alternative nitrogen source can release ammonium, which can be used subsequently by the active GS to synthesize Gln. Subsequently, the activation of NagB only by the uridylylated PII promotes the efficient utilization of amino sugars (Rodionova et al. 2018).

E. Research questions and objectives

As discussed previously, PII proteins are among the most widely distributed signaling proteins in nature and found in all domains of life. Canonical PII proteins bind to different effector molecules (ATP, ADP, 2-OG), the binding induces conformational changes, and enables the PII protein to bind to different targets to modulate different cellular functions. All cyanobacteria analyzed to date contain at least one *glnB* gene encoding for PII homologues. About 20% of sequenced cyanobacterial genomes were found to contain a second clear *glnB* paralogue. In addition to *glnB* paralogues, a close examination of available cyanobacteria genomes revealed further genes with similarity to *glnB* but lacking PII signature sequences, the putative gene products were termed as PII-like proteins. Notably, the functions for most of PII-like proteins are still unknown. Thus, it is tempting to speculate that PII-like proteins are involved in regulation of different cellular functions, which differ markedly from canonical PII proteins (Forchhammer 2010, Forchhammer & Lüddecke 2016).

In this study, we revised the PII superfamily to explore the relationships between different protein subfamilies within the PII superfamily and to assess the evolutionary distribution of proteins of the PII superfamily across different domains of life. To gain deep insights about the possible cellular functions of PII superfamily proteins, we chose members of three different subclasses (namely: 1- canonical PII-subfamily, 2- PII-like SbtB-subfamily and 3- PII-like CutA-subfamily) of the large PII superfamily and characterized them functionally, biochemically, and structurally. We used the advances of reverse genomics, proteomics, various biochemical/biophysical, and structural biology approaches to characterize the different subclasses of PII and PII-like proteins from cyanobacteria, algae (red and green) and land plants.

To address systematically the research questions, the research objectives of this work were summarized as following:

1- Structural and functional characterization of cyanobacterial PII-like protein SbtB.
2- Structural and functional characterization of cyanobacterial PII-like protein CutA.
3- Characterization of cyanobacterial PII residues crucial for N-acetyl-L-glutamate kinase (NAGK) and PipX interactions by site directed mutagenesis.
4- Characterization the sensory properties of canonical PII protein from the red algae in comparison to PII proteins from different phyla of oxygenic phototrophs to assess evolutionary conservation versus adaptive properties of PII proteins.
5- Characterization of metabolic adaptation strategies of the nonphotosynthetic alga *Polytomella parva*, and its consequence on the PII-mediated regulation of NAGK activity.

F. Results

1. Widespread distribution of the PII superfamily

The PII superfamily consists of widespread signal transduction proteins found in all domains of life. In addition to canonical PII proteins involved in C/N sensing, structurally similar PII-like proteins evolved to fulfill diverse, yet poorly understood cellular roles (Forchhammer & Lüddecke 2016).

To systematically explore the evolutionary relationship between different protein families that adopt the PII-ferredoxin fold and to build a comprehensive map of the PII-superfamily, with the help of Dr. Vikram Alva, we gathered homologs of all PII-like proteins of known structure and clustered them based on their all-against-all pairwise sequence similarity in CLANS. The obtained sequences were pooled together, filtered down to a maximum pairwise sequence identity of 60%, and clustered in CLANS (Fig. 2 in manuscript 1).

In the center of the obtained map, the canonical PII and SbtB proteins dominated two tightly connected clusters, from which all the other subgroups of the PII superfamily radiate (Fig. 2 in manuscript 1). At stringent P-value cut-offs, the members of canonical PII cluster remain connected to each, implying that the *glnB*, *glnk* and *glnZ* genes (encoding for canonical PII proteins), *nifI*, and the previously annotated PII-NG (encoding for PII proteins in genetic association with heavy metal efflux pump; HmeP) (Sant'Anna et al. 2009), are closely related and evolved together. However, for the SbtB cluster, at P-value cut-offs of 10e-29, the cluster divided into three distinct subclusters. The first subgroup refers to canonical SbtB proteins, which are in association with the bicarbonate transporter SbtA (Shibata et al. 2002, Du et al. 2014). The second group refers to a clade, which is associated with the carboxysome operon, the members of this subgroup are named carboxysomes associated PII proteins (CPII) (Wheatley et al. 2016). The members of the third subcluster are found in association with a SbtA-like transporter (Shibata et al. 2002), therefore we named it a SbtB-like subgroup. From the central PII/SbtB clusters, four extra PII-like clusters arise, referring to 1) DUF190, 2) DUF3240, 3) c-di-AMP receptor proteins PstA, and 4) multi-domain NIF3 proteins (Saikatendu et al. 2006, Godsey et al. 2007, Gundlach et al. 2015, Müller et al. 2015). In addition to the central clusters, further two distant clusters, including proteins of the divalent cation tolerance protein CutA and DUF2179, remain connected to the center of the map via NIF3 and PstA clusters, respectively. At the periphery of the map, at a P-value cutoff 1e-10, five groups represent the apparently most distant homologs, as they do not make any connections to the central of the map cluster. These groups comprise DUF2007,

two loosely connected clusters of proteases (Rhomboid_N and NRho), N-acetylmuramoyl-L-alanine amidase (AmiC), and HisG (multi-domain ATP-phosphoribosyl-transferase; the first enzyme of histidine biosynthesis). Based on our bioinformatic analysis we chose members of PII, SbtB, and CutA clusters to study them biochemically, physiologically, and structurally.

2. PII-like proteins in cyanobacteria

All cyanobacteria contain at least one *glnB* gene encoding a PII homologue. A close examination of cyanobacteria genomes available on CyanoBase (http://genome.microbedb.jp/cyanobase/) revealed further genes with similarity to *glnB* but lacking PII signature sequences, the putative gene products of such genes are referred to the PII-like proteins (referring to: SbtB, SbtB-like, and CutA subclusters). Hence, it would be important to find out to which signals PII-like proteins respond, and which targets they regulate for a better understanding of their cellular function, metabolic responses, and the genetic machinery system of different classes of PII-like proteins. Here, we studied the PII-like proteins SbtB and CutA from the cyanobacterial model organisms *Synechocystis* sp. PCC 6803, *Synechococcus elongatus* PCC 7942, and *Nostoc* sp. PCC 7120.

2.1. PII-like signaling protein SbtB is a conserved inorganic carbon sensor in cyanobacteria via integrating cAMP/AMP signals

Rubisco is the key enzyme for CO_2 fixation, however it is unable to discriminate between CO_2 and O_2, and so it is wasting energy through unnecessary O_2 fixation. With the evolutionary pressure of dropping atmospheric CO_2 levels, cyanobacteria have evolved a very efficient system, the carbon concentrating mechanism (CCM), to rise up the inorganic carbon (C_i) levels at Rubisco-carboxylating sites to ensure efficient photosynthesis (Burnap et al. 2015). The well characterized CCM-sodium-dependent bicarbonate transporter SbtA is highly expressed under C_i limitation together with the conserved uncharacterized PII-like SbtB protein (Wang et al. 2004, Schwarz et al. 2011, Du et al. 2014, Burnap et al. 2015). We were able to show physiologically that SbtB is an important component of the CCM in *Synechocystis* sp. PCC 6803 for efficient C_i acclimatization. The Δ*sbtB* mutant (ORF: *slr1513*) was unable to sense properly the availability of C_i levels (Figs. 3 & S7 in publication 2). The Δ*sbtB* mutant was unable to distinguish and switch easily between high/low carbon regimes and had a growth defect, especially after shifting from low carbon to high carbon conditions (Figs. 3 & S7 in publication 2). We were able to show as

well that SbtB localizes to the membrane under low carbon conditions, probably to regulate the SbtA uptake activity. The membrane localization of SbtB was strictly dependent on SbtA, as in a *sbtA* knockout mutant the sbtB protein was localized exclusively cytoplasmic (Fig. 2 in publication 2).

Biochemically, we were able to show that the SbtB protein is trimeric in solution and can bind in an anticooperative manner to a variety of adenine nucleotides (ATP, ADP, AMP, and cAMP) (Table 1 & Figs. S1 & S2 in publication 2). Structurally, SbtB showed the typical trimeric ferredoxin-like fold of the canonical PII proteins (Figs. 1 & S3 in publication 2). In addition, we were able to solve the crystal structure of SbtB in complex with AMP and cAMP and thereby identified the nucleotide-binding pocket (Figs. 1 & S4 in publication 2). The nucleotide-binding pocket was located between the lateral clefts of the SbtB subunit, perfectly matching the structure of the binding pocket of canonical PII proteins (Figs. 1 in publication 2), although the residues forming the binding pocket are poorly conserved.

Finally, we were able to provide the molecular basis for the recognition of cAMP as a novel C_i signaling molecule via binding to SbtB. Under low carbon conditions, the intracellular cAMP levels drop, while AMP levels increase (Fig. 4 in publication 2), implying that cAMP represents the high carbon signal and AMP is the low carbon signal. Constant with that, AMP was able to support the complex formation between SbtA-B proteins, while cAMP did not (Fig. 2 in publication 2). In addition, in the adenylyl cyclase deficient mutant (Δ*cya1*), which has a defect in cAMP production, the SbtB protein remained membrane localized in association with SbtA after the shift to high carbon conditions (Fig. 2 in publication 2). Since the intracellular level of cAMP is quite low in comparison to the AMP level, we performed a competition binding assay between the C_i signals cAMP/AMP. cAMP was able to replace the AMP-bound to SbtB under any tested condition, while AMP did not replace cAMP (Fig. 4 in publication 2), suggesting that under high carbon with elevated levels of cAMP and dropping AMP levels, cAMP would compete effectively to saturate SbtB. Therefore, our results suggest that SbtB, here described for the first time as a cAMP receptor protein of the PII superfamily, acts as a C_i sensor protein via integrating the energy state of the cell and cAMP binding, highlighting an evolutionary conserved role of cAMP/AMP as an indicator of the cellular carbon status across different domains of the life (Chen et al. 2000, Steegborn et al. 2005, Agostoni & Montgomery 2014, Burnap et al. 2015, Hennon et al. 2015).

2.2. PII-like CutA proteins do not confer heavy metal tolerance in cyanobacteria

Another PII-like protein with the lowest sequence identity to canonical PII proteins (Fig. 2 in manuscript 1) (Forchhammer & Lüddecke 2016), is the divalent ion tolerance protein CutA. Cluster mapping analysis revealed that the cluster of CutA proteins is evolutionary quite conserved and is universally distributed in all domains of the life (Fig. 1 in manuscript 1). The bioinformatic analysis suggested that the proteins of the CutA family were established in the last universal common ancestor and have remained largely conversed during the course of evolution. Since the exact cellular function of CutA proteins is still unknown, we chose members of cyanobacterial CutA proteins to study them physiologically, biochemically and structurally. Similar to canonical PII proteins, cyanobacterial CutA proteins are trimeric in solution and in crystal structure (Figs. 3 & S2 in manuscript 1), however, we were unable to identify the small signaling molecules binding to CutA protein. Nevertheless, the cavity between the inter-subunits of CutA proteins, the potential binding pocket, is formed by a number of conserved aromatic and charged residues (Figs. 4 & S5 in manuscript 1), suggesting a conserved binding/signaling function for CutA proteins. Next, to examine the possibility of the involvement of cyanobacterial CutA proteins in heavy metal tolerance, we created knockout and knockdown mutants in the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942, and the filamentous cyanobacterium *Nostoc* sp. PCC 7120, respectively. Unexpectedly, we were unable to link the cyanobacterial CutA to heavy metal tolerance. We found no difference between the Δ*cutA* mutants and wild-type strains growing with either no heavy metal addition or with different concentrations of metals (Cu^{2+}, Fe^{2+}, Pb^{2+}, Co^{2+}, Zn^{2+}, Cr^{2+}, Cd^{2+}, Ni^{2+}, Mn^{2+}) (Fig. 6 in manuscript 1), implying that CutA could play a different role in cyanobacteria rather than participating in heavy metal resistance.

3. Canonical PII proteins in cyanobacteria and Eukaryotes

Canonical PII proteins are involved in the regulation of various carbon and nitrogen metabolism related processes (Huergo et al. 2013, Forchhammer & Lüddecke 2016). For this purpose, prokaryotic PII proteins act as an energy/carbon/nitrogen sensor via binding of ATP, ADP and the TCA cycle intermediate 2-OG, which enable PII proteins to integrate different signals to regulate different targets according to the metabolic situation. Although the canonical PII proteins are well studied in prokaryotes (Huergo et al. 2013), the eukaryotic PII proteins are still poorly investigated.

Therefore, we focused on eukaryotic PII proteins in comparison to cyanobacterial PII proteins, as the cyanobacterial PII proteins are the most closely related to eukaryotic PII proteins, which originated from the cyanobacterial endosymbiotic ancestor of the plant kingdom.

3.1. Key amino acid residues of cyanobacterial PII protein required for NAGK and PipX interactions

In cyanobacteria, PII regulates the nitrogen metabolism by interacting with: 1) N-acetyl-L-glutamate kinase (NAGK), the key enzyme of arginine biosynthesis, and 2) PII-interacting protein-X (PipX), the co-activator of the global cyanobacterial nitrogen-transcription factor NtcA. The binding of PII to NAGK increases the NAG/NAG-P turnover and prevents arginine-feedback inhibition to NAGK. The binding of PII to the PipX protein leads to a deactivation of the NtcA regulon. To figure out the key residues for active cellular function of PII protein, we created different point mutations in the cyanobaterial *Synechococcus elongatus* PCC 7942 PII protein (*Se*PII). We used a variety of *in vitro* approaches including ITC, SPR, SEC-MALS and enzymatic based assays to study the influence of K58N and I86N point mutations on the PII mediated NAGK and PipX interactions. We were able to show that the Lys58 is a key residue for the sensory properties of *Se*PII and for mediating PII interactions. The K58N mutation causes the loss of 2-OG binding (Fig. S1 in publication 5) and has a strong negative effect on ADP (Fig. 1 in publication 5), NAGK (Figs. 2 & S2 in publication 5) and PipX (Fig. 4 in publication 5) binding. The replacement of the nearby Leu56 by Lys can partially compensate the loss of K58 and enhanced the binding of *Se*PII (K58N/L56K) variant towards ADP (Fig. 1 in publication 5), NAGK (Figs. 2 & S2 in publication 5) and PipX (Fig. 4 in publication 5).

Previously, it was demonstrated that the *Se*PII (I86N) variant was able to constitutively activate NAGK *in vitro* and *in vivo*, leading to an accumulation of arginine and cyanophycin (Fokina et al. 2010b and Watzer et al. 2015). To gain deeper insights in the signaling mechanism of the *Se*PII (I86N) variant with respect to PipX interaction, Björn Watzer investigated the *in vitro Se*PII (I86N)-PipX interaction using SPR, revealing that the *Se*PII (I86N) variant forms a stable complex with PipX in presence of ATP and even without effectors molecules (i.e. without ADP stimulation). However, 2-OG did not have strong inhibitory effect on the PII (I86N)-PipX complex (Fig. 5 in publication 5).

During purification of the recombinant PII variants, we observed a strong growth defect of the competent LEMO21 (DE3) *E. coli* strain harboring the plasmid encoding for either PII (K58N) or (K58N/L56K) compared to PII (WT) directly after the induction. To study this phenomenon

systematically, Michael Haffner (a master student in our department) performed further experiments under my supervision. First, to rule out the possibility that the formation of heterotrimers between *E. coli* PII and SePII (Forchhammer et al. 1999) is the cause of the growth defect, the previous growth experiments were repeated in a ΔPII deletion background (*E. coli* RB9060 strain) (Bueno et al. 1985). With expressing PII (K58N) and (K58N/L56K) variants in the *E. coli* RB9060 strain, we observed a similar growth defect (data not shown), excluding the PII heterotrimer formation to be reason for the observed growth disadvantage. To deeply investigate what could be the cause of the growth defect, we performed metabolomic (in collaboration with Dr. Klaus Brilisauer) and pulldown-based proteomic (samples were submitted to Tübingen Proteome Center) analysis. In metabolomic analysis, we observed an accumulation of acetylspermidine and diacetylspermidine compounds in *E. coli* cells expressing PII (K58N) and (K58N/L56K) variants, while the cells expressing the PII (WT) protein accumulated higher levels of the triglyceride compounds. In the pulldown-based assay, we identified several proteins pulled with the PII (WT, K58N, and K58N/L56K) proteins. Notably, we identified several targets involved in fatty acid biosynthesis (e.g. BCCP subunit of ACCase, acyl carrier protein, 3-hydroxyacyl-[acyl-carrier-protein] dehydratase). These results imply that the PII variant could potentially influence the fatty acid metabolism.

3.2. Coevolution of hetero-oligomeric PII-NAGK enzyme in a non-photosynthetic alga *Polytomella parva*

In the course of evolution, several algae and plants became heterotrophic by losing their photosynthetic ability, however, in most cases, a non-photosynthetic plastid was maintained to fulfil various metabolic functions that are not related to photosynthesis (Molina et al. 2014, Smith & Lee 2014). Among these organisms, the colorless alga *Polytomella parva* is a special case, as its plastid is devoid of any DNA, but the organelle is maintained for specific metabolic tasks carried out by nuclear-encoded enzymes. *Polytomella* is believed to originate from a green ancestor closely related to the photosynthetic alga *Chlamydomonas reinhardtii* by the loss of the entire chloroplast genome and the concomitant loss of the photosynthetic apparatus (Smith et al. 2013, Smith & Asmail 2014). This makes *P. parva* an attractive model system for exploring molecular evolution events in response to fundamental metabolic alterations imposed by a life-style switch from autotrophic to heterotrophic lifestyle by the loss of photosynthesis. However, the metabolic adaptation strategies of *P. parva* remained elusive so far. Therefore, we characterized the

metabolic adaptation strategies of *P. parva* in response to nitrogen starvation, in comparison to the closely related photosynthetic alga *Chlamydomonas reinhardtii*.

Metabolomic analysis revealed that *P. parva* accumulates significantly higher amounts of most TCA and GS/GOGAT cycles intermediates as well as glutamate, aspartate and arginine, the latter being specific for the colorless plastid (Fig. 1 in publication 3). To explain mechanistically, how this altered metabolite status affects metabolite sensing in *P. parva*, we studied next as a test case the PII-mediated regulation of NAGK activity, which controls the committed step of arginine biosynthesis. In the plant kingdom (except for *Brassicaceae*), PII controls NAGK activity in a glutamine-dependent manner, through the evolution of a C-terminal extension, called the Q-loop, which represents the Gln binding site (Chellamuthu et al. 2014). Similarly, we found that *P. parva* PII enhances NAGK activity in response to glutamine levels (Figs. 4 & 5 in publication 3), despite a stable complex being formed in a glutamine-independent manner (Figs. 3 & 6 in publication 3). So, we hypothesized that the PII signaling protein, so far known to transiently interact with NAGK, emerges as a subunit of the enzyme NAGK. Using SPR and SEC-MALS analysis, we provided evidences that *P. parva* PII forms an extremely stable complex with NAGK independently of any effector molecules (Figs. 3, 6 & S4 in publication 3). Without PII, the NAGK protein isn't able to properly sense the feedback inhibitor arginine. In this hetero-oligomeric complex, PII tunes the response of NAGK towards the feedback-inhibitor arginine in a glutamine-dependent manner (Figs. 4 & 5 in publication 3). In absence of Gln and under elevated levels of arginine, *P. parva* PII inhibits NAGK (Fig. 4 in publication 3). Moreover, the metabolites 2-OG and ATP/ADP levels do not affect PII regulation of NAGK (Figs. 5 & 6 in publication 3), indicating that the *P. parva* PII-NAGK system has lost the ability to estimate the current energy and carbon status of the cells but specialized to provide an entirely glutamine-dependent arginine-feedback control, highlighting the evolutionary plasticity of the PII signaling system. Furthermore, heterologous experiments using PII and NAGK proteins from *Chlamydomonas* and *Oryza sativa* (rice) demonstrated that the ability to form the permanent PII-NAGK complex in *P. parva* is intrinsic to both PII and NAGK proteins (Figs. 4, 6, 7, S3 & S4 in publication 3), which is a unique example for the evolution of enzyme complexes.

3.3. Sensory properties of canonical PII proteins in Eukaryotes versus Cyanobacteria

3.3.1. PII signaling transduction system in Rhodophyta represents an evolutionary intermediate between Chloroplastida and cyanobacterial systems

From our previous study, we demonstrated that the PII protein from *P. parva* lost the ability to sense ADP and 2-OG. Therefore, we wanted to expand our analysis to assess the sensory properties of other representative members of the eukaryotes. Eukaryotic PII homologues are restricted only to the kingdom of plantae. Eukaryotic PII homologues have been identified from Chloroplastida (green algae and land plants), where they are nuclear-encoded, and in Rhodophyta (red algae), where they are coded by the plastid genome (Uhrig et al. 2009). In both groups, PII is localized in the chloroplast (Hsieh et al. 1998, Ermilova et al. 2013). Notably, in Chloroplastida, PII proteins control the ornithine pathway via regulation of NAGK, as in cyanobacteria (Mizuno et al. 2007, Beez et al. 2009, Chellamuthu et al. 2014). In addition, the signaling roles of PII proteins are well-studied in Chloroplastida, Bacteria and Archaea (Beez et al. 2009, Fokina et al. 2010b, Chellamuthu et al. 2014, Palanca et al. 2014), but no PII homolog has been characterized so far from the red algae.

In this work, we were able to provide the first biochemical characterization of red algal PII protein from the model organism *Porphyra purpurea*. Moreover, we comparatively characterized the sensory features of PII proteins across different phyla of oxygenic photoautotrophs (prokaryotic cyanobacteria, eukaryotic: green alga *Chlamydomonas* and higher plant *Physcomitrella*) to address the evolutionary sensing properties of PII proteins. Like its cyanobacterial PII homologs (Fokina et al. 2010b & 2011), the red algal PII protein binds competitively ATP/ADP in anticooperative manner and 2-OG in synergy with ATP (Table 1 & Fig. 2 in publication 1). Surprisingly, the land plant PII protein was unable to bind ADP and 2-OG (Table 2 & Fig. 5 in publication 1), however the green alga PII protein was not able to bind ADP only (Table 2 & Fig. 4 in publication 1). Interestingly, we were unable to detect binding for ATP for *Chlamydomonas* PII protein unless 2-OG is present, implying that the affinity towards ATP is low, while in presence of 2-OG the ATP affinity increases dramatically (Table 2 & Fig. 4 in publication 1). For the red algal and cyanobacterial PII proteins, similarly, 2-OG enhanced the affinity of ATP binding (Table

1 & Figs. 2 & 3 in publication 1). However, unexpectedly 2-OG negatively affected ADP binding in the presence of Mg^{2+}, for the red algal and cyanobacterial PII proteins (Table 1 & Figs. 2 & 3 in publication 1). In contrast to eukaryotic PII proteins, the red alga PII protein enhanced the kinetics of NAGK and relieved it from the Arg feedback inhibition in a Gln-independent manner (Figs. 6 & 7 in publication 1). Analogous to cyanobacterial PII proteins, the NAGK activation by the red algal PII protein was partially suppressed by 2-OG (i.e. 2-OG induces the PII-NAGK complex dissociation) and ADP (i.e. due to the competition between ADP and ATP for the available PII binding sites) (Figs. 7 & 8 in publication 1). Despite high sequence identity between the red algal and cyanobacterial PII proteins (Fig. 1 in publication 1), the red algal PII protein was not able to interact with the cyanobacterial transcriptional activator PipX (Fig. S3 in publication 1).

3.3.2. Cyanobacterial PII proteins do not sense Ca^{2+} ions

In cyanobacteria, a transient increase in the intracellular Ca^{2+} levels was observed in *S. elongatus* PCC 7942, in response to nitrogen limitation. *In vivo* studies revealed that the PII and the NtcA proteins are required to trigger the transient Ca^{2+} signal (Leganés et al. 2009). In addition, the Ca^{2+} ions were suggested to influence the PII modification (Zhao et al. 2005), however this hypothesis remains experimentally untested. Therefore, it was speculated that PII proteins can sense metals ions, in particular Ca^{2+} ions, as several crystal structures deposit in the protein data bank were described to bind to bind divalent cations, for examples PII proteins from: *Synechocystis* sp. PCC 6803 was found to bind Ca^{2+} (PDB: 1UL3), *Synechococcus elongatus* sp. PCC 7942 was found to bind Ni^{2+} (PDD: 1QY7), *E. coli* was found to bind Mg^{2+} (PDB: 5L9N), *B. subtilis* was found to bind Zn^{2+} (PDB: 4R25), *C. reinhardtii* was found to bind Cd^{2+} (Xu et al. 2003, Chellamuthu 2013). Thus, it seems that Ca^{2+} influences the sensing of nitrogen availability, possibly through PII sensing machinery.

To address this question experimentally, we tested the ability of the recombinant PII protein from *S. elongatus* PCC 7942 to bind Ca^{2+} *in vitro* using ITC. After excessive dialysis of the PII protein in Chelex 100 to get rid of any divalent cation contamination as described in Su et al. (2011), the ITC experiments revealed that PII protein was not able to bind Ca^{2+} (Fig. 17). Nevertheless, we wanted next to investigate the influence of fluctuation of extracellular Ca^{2+} concentrations by addition of 1 mM Ca^{2+} ($BG11_{+Ca}$) or remove of Ca^{2+} completely ($BG11_{-Ca}$) from the BG11-media on the growth of Δ*PII* mutant in comparison to the wildtype stain. Since the Δ*PII* mutant previously showed a growth disadvantage compared to the wildtype (Fig. 17), we normalized the efficiency of the photosystem II (PSII) and the rate of the O_2 evolution to the untreated cells (taken as 100

%). Even though, under all tested conditions the Δ*PII* mutant behaved similarly to the wildtype (Fig. 17) with respect to the PSII efficiency, the rate of the O_2 evolution, and the pigmentation. This behavior implies that the PII protein does not sense Ca^{2+} ions directly.

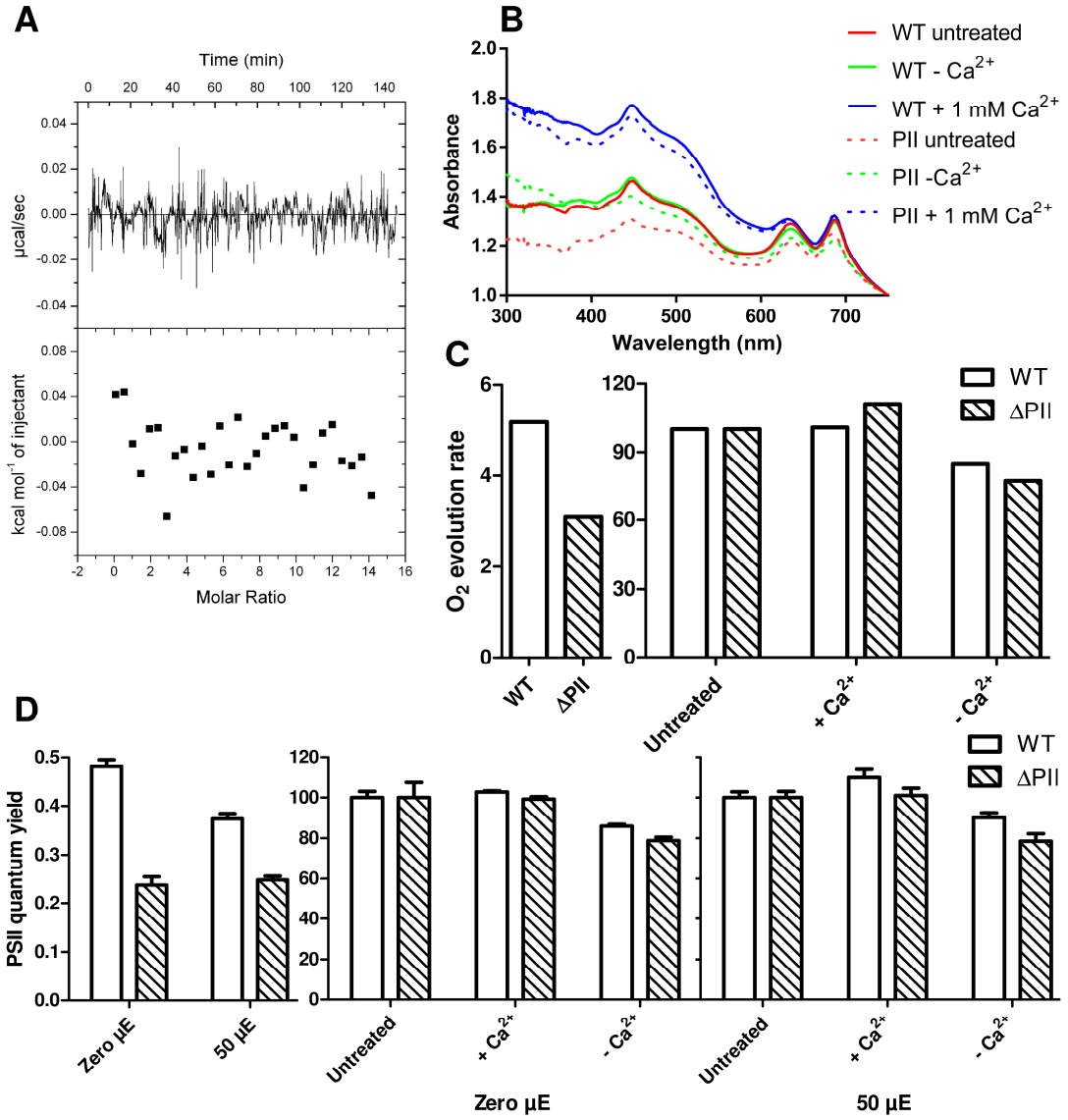


Fig. 17. Cyanobacterial PII proteins do not sense Ca^{2+} ions. (A) ITC analysis for binding of Ca^{2+} to the recombinant PII protein revealed no signal upon titration of Ca^{2+} ions. (B) Absorption spectra for wildtype (WT) and Δ*PII* strains after 6 days of Ca^{2+} treatment, as indicated. (C) The rate of the O_2 evolution ($nmol.ml^{-1}.min^{-1}$) after 65 h of Ca^{2+} treatment. (D) Efficiency of photosystem II (PSII) quantum yield at 0 and 50 µE after 65 h of treatment Ca^{2+} treatment. The PSII quantum yield and the rate of O_2 evolution were normalized to the untreated cells (taken as 100 %), the difference between WT and Δ*PII* strains for untreated cells was indicated in the right panel. B, C and D revealed no significance difference between WT and Δ*PII* upon addition of 1 mM or removal of Ca^{2+} ions completely from the medium; the untreated cells were used as standard for comparison.

4. A novel cyanobacterial Ca^{2+}-binding protein influences the photosynthetic activity

Ca^{2+} was speculated to be among the most important second messengers in cyanobacteria (Agostoni & Montgomery 2014). However, it is poorly understood how cyanobacteria sense Ca^{2+}. To identify a potential Ca^{2+} sensor protein, the searches for Ca^{2+}-binding domains in cyanobacterial genome databases revealed a conserved gene encoding a putative protein with two characteristic Ca^{2+}-binding EF-hand domains in filamentous cyanobacteria (Fig. 1 in publication 4). To investigate the possible signaling function of the putative Ca^{2+} sensor EF-hand (CSE) protein, the recombinant CSE protein encoded by *asr1131* from cyanobacterium *Nostoc* sp. PCC 7120 was overexpressed with N-terminal His-tag. In solution, the CSE protein is monomeric and exhibit a typical α-helical profile (Fig. 2 in publication 4). With biochemical and biophysical approaches, we were able to show that the CSE protein can bind Ca^{2+} and Mg^{2+} *in vitro* (Table 1 & Figs. 2 & S2 in publication 4). Competition binding assays revealed that CSE protein can bind Ca^{2+} even under elevated levels of Mg^{2+} (2 mM). In contrast, under saturating conditions of Ca^{2+} (150 µM), the Mg^{2+} binding to the CSE protein was completely suppressed (Table 1 & Fig. S2 in publication 4), implying that the CSE protein would bind Ca^{2+} specifically under physiological conditions. To identify the cellular function of the CSE protein, collaborator partners at Turku University performed several physiological experiments with knockout and overexpressing mutants in *Nostoc* sp. PCC 7120. The overexpression mutant showed growth disadvantage and reduced photosynthetic activity compared to the wild-type strain (Figs. 3 & 5 in publication 4). Transcriptional analysis indicated a strong downregulation of genes involved in the assembly of functional phycobilisome (PBS) complexes (Table 3 in Sub. manuscript 2). By means of biophysical and biochemical approaches, we were able to show that the energy transfer between PBS and the photosystems is impaired in the mutant, leading to the disruption of PBS-PSI supercomplexes and PSII dimers (Figs. 4-8 in publication 4).

G. Discussion

1. Distribution of PII superfamily

It appears that the PII superfamily is more widely distributed than it has been ever thought. The cluster analysis of PII superfamily revealed that the PII, CutA, HisG and Nif3 clusters are widely distributed within prokaryotes (bacteria and archaea) and eukaryotes, while the other clusters are found in prokaryotes (bacteria and archaea) only. Out of the restricted prokaryotes clusters, three of them (SbtB, DUF190 and DUF2179) are distributed in bacterial and archaea, while the other clusters are restricted to bacterial origin only (Fig. S1 in manuscript 1). Although the eukaryotic-PII proteins have only been evolved in Chloroplastida via endosymbiosis, the CutA proteins were found ubiquitously distributed from prokaryotic to mammalian cells. The other members of the PII superfamily presumably were linearly inherited by bacteria and archaea, or alternatively they arose later in the course of the evolution by gene duplication events for rapid functional adaptation. This suggest that CutA protein may have been present in the last universal common ancestor.

Notably, previous phylogenetic analysis pointed out two types of SbtA in cyanobacteria (Shibata et al. 2002), one refers to the verified bicarbonate SbtA transporter, while the other member refers to a low sequence homolog called SbtA-like with uncharacterized uptake activities. Interestingly, *SbtAB*-like operon exists in some cyanobacterial strains along with a SbtAB operon, e.g. *Anabaena* sp. PCC7120 possesses both types of *SbtAB/like* operons. This arrangement indicates that *SbtB and new subgroup of SbtB-like* genes are products of an early gene duplication event. Remarkably, the tight connection of the DUF190 cluster to the SbtB cluster, and the DUF3240 to the PII cluster imply that they could play a role in inorganic carbon and nitrogen metabolism, respectively.

2. SbtB a new inorganic carbon sensor protein in cyanobacteria

With the identification the signaling role of SbtB, we describe in this work a discovery of major importance: a functional link between the second messenger cyclic-AMP signaling and CO_2/bicarbonate assimilation. Although it was known for a while that a wide range of soluble adenylate cyclases respond to CO_2 availability (Cann et al. 2003, Steegborn et al. 2005, Hammer et al. 2006, Raven 2006), the functional connection between cAMP and CO_2/bicarbonate

metabolism remained elusive. In our work, we presented SbtB as a novel senor protein that responds to cAMP and is involved in bicarbonate/CO_2 uptake regulation. We found that the SbtB protein is ubiquitously distributed in cyanobacteria, which are responsible for at least 10% of todays global CO_2 fixation. There is currently interest in introducing the cyanobacterial CCM into crop plants to increase the efficiency of photosynthesis (Long et al. 2016). Therefore, our understanding of SbtB as a component of the CCM could help in future to improve the plant photosynthetic efficiency, in addition to a better understanding of the important ecological role of cyanobacteria.

The cumulative evidences imply that cyclic AMP signaling plays a central role in inorganic carbon perception. Our results indicated that cAMP appears to represent a high C_i signal, in agreement with the finding that the activity of evolutionary conserved cyanobacterial soluble adenylyl cyclase was stimulated under elevated levels of C_i and led to increased cAMP production (Cann et al. 2003, Steegborn et al. 2005, Hammer et al. 2006, Raven 2006). Additionally, it was reported that the elevated CO_2 levels in diatoms enhanced the cAMP production by sAC, which in turn promotes down regulation of the algal CCM (Hennon et al. 2015). Thus, cyclic AMP can be seen as an evolutionary conserved signaling molecule responding to C_i fluctuation and regulating the CCM in distantly unrelated phyla of algae and cyanobacteria.

Structurally, although the residues forming the nucleotide binding pocket of SbtB are only poorly conserved in comparison to canonical PII, the position of bound nucleotides matches perfectly with the canonical PII proteins (Fig. 1 in publication 2). This clearly indicates that the proteins of the PII superfamily, in particular of SbtB and canonical PII clusters (Fig. 2 in manuscript 1), arose evolutionary from a common ancestor, whose nucleotide binding pocket has been architecturally preserved. However, the binding mode has evolved with few mutations in the residues forming the binding pocket to sense different adenyl nucleotides or, in combination with other effectors (e.g. ATP-dependent binding of 2-OG to canonical PII proteins), to fulfill different signaling tasks. Notably, from the SbtB cluster (Fig. 2 in manuscript 1), a member of the CPII subgroup has been recently described from the chemoautotrophic bacterium *Thiomonas intermedia* to sense AMP and ADP (PDB: 5DS7), in addition it exerts its regulatory function through binding to HCO_3^- (Wheatley et al. 2016). Interestingly, the proteins of the SbtB and CPII subclusters bind ADP and AMP. While the CPII protein responds directly to HCO_3^- and thereby increases the affinity to the adenyl-nucleotides, the SbtB protein does not sense HCO_3^- but instead binds to the second messenger cAMP, which reflects the carbon status of the cell. Since both proteins are involved in the regulation of C_i assimilation and have in common the response to ADP/AMP, it seems that the main task of these proteins is linking the C_i assimilation reactions with the energy state of the

cell. Although the sensing of adenyl-nucleotides is conserved, the sensing of the C_i status occurs via distinct mechanisms.

Collectively, our results support a model whereby the SbtB protein senses AMP as a low carbon signal to trigger the membrane localization of SbtB to regulate SbtA under low carbon conditions. However, under high carbon conditions, the elevated levels of cAMP compete with AMP for the available SbtB binding sites, which in turn induces SbtB dissociation from the membrane, thereby disrupting the SbtA-B complex (Fig. 18). Hence, SbtB represents a unique mechanism of cAMP mediated signaling in C_i assimilation, confirming the conserved signaling task of the secondary messenger cAMP in the CCM, and thereby this study provides novel insights into cell signaling in prokaryotes.

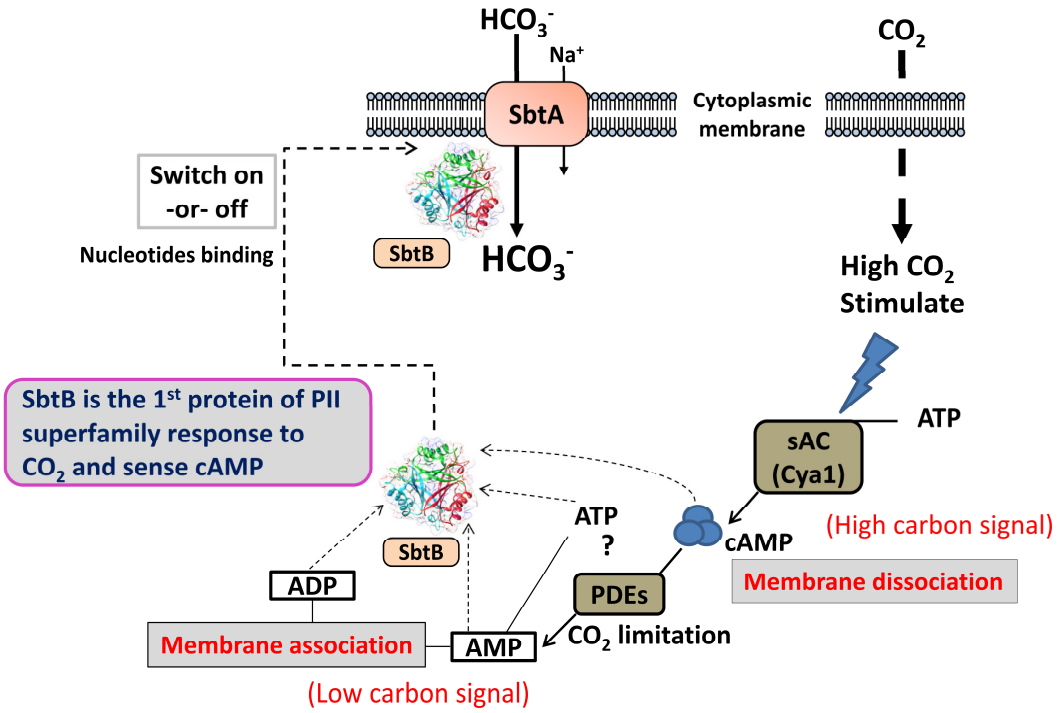


Fig. 18. Model for the proposed regulatory mechanism of the PII-like protein SbtB.

3. The PII-like protein CutA in cyanobacteria

For the PII-like protein CutA, the structure corresponds to the trimeric ferredoxin-like assembly of canonical PII proteins. Despite the low sequence conservation between CutA, SbtB, and canonical PII proteins, the superimposes of CutA over SbtB and PII revealed that they perfectly overlaid with less than 1.0 Å rmsd (Figs. 4 & S4 in manuscript 1), highlighting evolutionary hallmark features of the trimeric core architecture of the PII superfamily. Hence, we concluded that the close structural similarity of the cell signaling PII/PII-like proteins supports again our

hypothesis that they have probably evolved early from an ancestral sensor protein that already adapted the characteristic trimeric core architecture.

To our surprise, we were unable to show the ability of cyanobacterial CutA proteins neither to bind Cu^{2+} *in vitro* nor to confer heavy metal tolerance *in vivo*. Knockout and knockdown mutants of *cutA* generated in two independent cyanobacterial strains, *Nostoc* sp. and *S. elongatus*, revealed that *cutA* genes could be involved in a different sensory function rather than metal sensing. Constant with our phenotypic observations, the gene expression of *cutA1* in the plant pathogen *Xylella fastidiosa* showed no induction or weak induction without significance changes under elevated levels of external Cu^{2+} concentrations in two different growth media (Rodrigues et al. 2008). Interestingly, also in eukaryotic plants it seems that the CutA protein is not involved in metal tolerance, as the *CutA* knockout lines in *Arabidiosis thaliana* revealed that CutA is not essential for Cu^{2+} tolerance (Burkhead et al. 2003). However, it's still possible that CutA plays a signaling or regulatory role related to heavy metals via the regulation of heavy metal efflux pumps similar to the other members of the PII superfamily, like *GlnK* and *SbtB* which regulate ammonium or bicarbonate transport channels, respectively (Radchenko et al. 2010). Thus, the CutA1 function may occur by ion binding (not likely) and/or by affecting ion import/export through interaction with membrane transporters.

Also, we cannot rule out the wrong annotation for involvement of the CutA protein in heavy metal tolerance because of the lack of physiological *in vivo* data about the role of CutA in prokaryotes. The initial study conducted by Fong et al. (1995) in *E. coli* showed that the *cutA* locus contains three genes, in two operons, encoding for a cytoplasmic cytoplasmatic protein of ~13 kDa (CutA1) in one operon and two inner membrane proteins (CutA2-3) in the second operon. The mutation of the *cutA* locus revealed sensitivity of *E. coli* cells to divalent metal ions. The complementation of Δ*cutA* locus with the ORF encoding for CutA1 conferred enhanced tolerance of the mutant only to elevated Cd^{2+} levels, while it did not confer tolerance against Cu^{2+}, Zn^{2+}, or Ni^{2+}. Instead, the complementation using the ORF encoding for both CutA1 and 2 restored the tolerance of the mutant toward high Cu^{2+} and Ni^{2+} levels. Therefore, they concluded that both operons are potentially involved in heavy metal tolerance. Since then, the scientific community believed in this finding without further physiological examination.

4. Cyanobacterial PII variants (K58N) and (I86N)

Previous structural analysis of PII proteins in various complexed states from various organisms indicated that Lys58 plays an important role in the function of PII. However, the K58 residue is not

involved in direct interactions with ADP or ATP, but it rather seems to be important to anchor the B-loop via a H-bond to residue G87 (PDBs: 2XUL, 3MHY, and 2XBP). The ADP binding events induce conformational changes within the surface exposed T-loop enabling K58 to interact with Q39. Therefore, K58 can serve as well to anchor the T-loop in the ADP bound conformations via a hydrogen bond to Q39 (PDB: 4CNZ), which characterizes/facilitates the preferred ADP bound state (Truan et al. 2014, Zeth et al. 2014). Further, it anchors also the T-loop in the bent conformation found in the PII-NAGK complex via a salt-bridge to E44 (PDBs: 2V5H & 2XBP) (Llácer et al. 2007, Truan et al. 2010 & 2014, Fokina et al. 2010a&b). Remarkably, the 2-OG binding is coordinated mainly by the highly conserved residues K58 and Q39 (Fokina et al. 2010a, Truan et al. 2010). Therefore, upon 2-OG binding, the K58-E44 interaction is broken and replaced by a new salt bridge between K58 and 2-OG (PDBs: 2XUL & 3MHY). Elegantly, the C5 carboxyl group of 2-OG quite precisely replaces the E44 carboxyl group, such that K58 does not need to change its conformation.

Our *in vitro* assays confirmed that the K58 is a key residual for proper signaling function of PII via affecting on: 1) sensory properties of ATP/ADP/2-OG and 2) the interactions with NAGK and PipX. The PII (K58N) variant is still able to bind ATP but lost the ability to bind 2-OG and strongly influenced ADP binding negatively. The loss of 2-OG sensing is expected, since the K58 contribute binds 2-OG directly. Constant with our results, previous results demonstrated that the PII (K58M) was not able to bind to 2-OG (Fokina et al. 2010a). However, we strongly believe that the defect in anchoring the T-loop (and maybe B-loop) is most likely the cause for the change of the affinities of the adenyl nucleotides (ATP and ADP). Thus, it seems that anchoring the T-loop is crucial for nucleotide binding. Apparently, the mutation of K58 destabilizes the PII T-loop structure, since the NAGK and PipX interactions which are mediated by the T-loop are influenced by the K58N mutation. In case of NAGK, the T-loop of the PII (K58N) variant can still insert into the NAGK to activate it, however in presence of arginine the PII (K58N)-NAGK complex became more sensitive and dissociated. Therefore, we concluded that the tight complex formation between PII and NAGK is required to relief NAGK from Arg feedback inhibition. For the PII (I86N) variant, the *in vitro* ability of the variant to form a strong complex with PipX suggests that sequestration of the PipX by PII (I86N) variant *in vivo* would tune down the NtcA regulon.

The biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase (ACCase) is a common target between cyanobacterial and *E. coli* PII protein (Gerhardt et al. 2015, Hauf et al. 2016b). The binding of PII to the BCCP subunit tunes down the ACCase activity. Indeed, we were able to pulldown the BCCP subunit of ACCase and other targets involved in fatty acid metabolism with PII (K58N, K58N/L56K, and WT) variants. Apparently, the PII variants may influence the *in*

vivo ACCase activity, leading to alternation of acetyl-CoA levels which can be then used by other pathways, explaining the accumulation of acetylspermidine and diacetylspermidine compounds. The growth defect of the *E. coli* could be explained by impairment of fatty acid metabolism, as the strain expressing PII (WT) protein accumulated higher levels of the triglyceride compounds. However, the direct influence of the PII variants on ACCase activity should be investigated further, e.g. by *in vitro* enzymatic assays.

Based on our observation in *E. coli*, it would be interesting to generate a cyanobacterial strain that harbors the PII (K58N) variant to investigate the influence of the change in the sensory properties on the physiology of cyanobacteria, possibly to generate a strain accumulating high levels of acetyl-CoA for industrial applications. In addition, it would be helpful as well to characterize the *in vivo* influence of PII (I86N) on the NtcA regulon (i.e. due to the strong sequestration of the NtcA coactivator PipX by the I86N variant) in response to nitrogen limitation conditions using transcriptome or proteome analysis, as the *Synechocystis* sp. BW86 strain carrying the PII (I86N) variant is available (Watzer et al. 2015).

5. Coevolution of hetero PII-NAGK complex in nonphotosynthetic alga *P. parva*

P. parva represents an attractive model system for exploring the metabolic adaptation strategies imposed by life-style switch from autotroph to heterotroph by the loss of photosynthesis. We were able to show that the evolution of non-phototrophic metabolism in *P. parva* alga involves a unique adaptation of NAG kinase and PII sensing. As a consequence of the purely heterotrophic lifestyle switch, the metabolomics analysis revealed that in *P. parva* the mitochondrial metabolites are dominating with huge accumulation of TCA cycle and GS/GOGAT intermediates, whereas the degenerated plastid appears to function as a highly efficient nitrogen assimilation compartment, with the accumulation of glutamate, aspartate and arginine.

Correlating with the altered metabolite status, the plastid-localized C/N sensory PII signaling protein and its regulatory-target NAGK shows unique adaptation features. The versatile PII signal transduction protein and NAGK have co-evolved towards an extremely stable hetero-oligomeric complex, independent of any effector molecules, with PII as a stably bound regulatory subunit of the enzyme. The very high stability of the complex suggests that *in vivo* these proteins probably are always complexed, irrespective of effector molecules. So, the PII signaling protein, which was so far known as a transiently interacting signaling protein, has evolved into a stably bound subunit of the NAGK enzyme. In this hetero-oligomeric complex, NAGK senses arginine, while PII senses

glutamine. The low amount of Gln necessary to activate NAGK fits with the enhanced production of Arg. We propose that the NAGK is always ''ON'' unless there is severe nitrogen (Glutamine) limitation. That means, the default mode of the system is high arginine production through high NAGK activity. Since the sequence identity between PII and NAGK proteins from *P. parva* and *C. reinhardtii* is high, it seems that a few amino acid substitutions were sufficient to change the transient PII-NAGK complex, as in all oxygenic phototrophs (Llácer et al. 2007; Chellamuthu et al. 2014). For future perspective, it would be interesting to determine the atomic coordinates of the PII-NAGK complex form *P. parva* using X-ray crystallography. Interestingly, it seems that the PII protein in *P. parva* lost the ability to estimate the current energy and C/N status of the cell by losing the ability to sense 2-OG and ADP. This suggests that the sensing properties of PII proteins is easily adjustable to the regulatory need of the respective metabolic situation in the individual organisms, highlighting the evolutionary plasticity of PII signaling.

6. Sensory properties of PII proteins and evolution of PII signaling system in Rhodophyta

We branded relevant sensory properties of PII proteins from different phyla of eukaryotic Rhodophyta and Chloroplastida (green algae and land plants) and prokaryotic Cyanobacteria to evaluate the evolutionary conservation versus adaptive features. The PII protein from the red alga *P. purpurea* was characterized with respect to its functional/signaling similarities and differences compared with PII proteins from oxygenic photoautotrophs of cyanobacteria and representative Chlorophyta plants (green algae and mosses). In particular, we were able to highlight the PII signaling transduction system in Rhodophyta as an intermediary evolutionary state between Chlorophyta and Cyanobacteria. Our data suggest that the sensory properties of the red algae PII resembles the cyanobacterial state very much (Fokina et al. 2010b & 2011), whereas in later phases of Viridiplantae evolution, the PII proteins deviated in their signaling features, becoming very heterogeneous with respect to ADP and 2-OG sensing, while only the binding of ATP has been conserved. Moreover, our results emphasize also a specific role of PipX as a partner only to cyanobacterial PII and confirm a conserved role the controlling enzyme of arginine biosynthesis, NAGK, as the primary PII target throughout the evolution of oxygenic phototrophs from cyanobacteria to Archaeplastida.

It seems that the anticooperative nature for the binding of the small effector molecules (ATP, ADP, and 2-OG) is a characteristic feature of almost all classical PII proteins (Smith et al. 2003, Fokina et al. 2010a, b, 2011, Helfmann et al. 2010, Ma et al. 2013, Oliveira et al. 2015). Previously, it

was reported that the plant PII protein from *Arabidopsis thaliana* was able to bind 2-OG in presence of ATP or ADP, but with less affinity in presence of ADP (Smith et al. 2003). However, in our study, no binding of 2-OG was detected for all tested PII proteins in the presence of ADP, in agreement with ATP-dependent binding of 2-OG for cyanobacterial PII (Fokina et al. 2011).
It's well-known that the fluctuation of the energy state of the cell (ATP/ADP ratio) depends on several conditions, including growth phases, nutrient availability, and carbon and nitrogen availability. Therefore, the physiological fluctuation of the ATP/ADP ratio allows PII proteins to response differentially to modulate different targets according to the current energy state of the cell. The results presented here, suggest that under nitrogen rich conditions with lower concentrations of 2-OG, there would be a competition between ATP and ADP for the occupation of the available binding sites of PII depending on the physiological concentrations of the nucleotides at that particular moment. Under poor nitrogen supply, the levels of 2-OG would rise, which in turn form a complex in the already occupied PII:ATP sites and for unoccupied PII sites, it would favor the ATP binding rather than ADP. Hence, the fluctuation of cellular 2-OG levels seems to be pivotal influential on the competition between ATP/ADP for the available sites of the PII proteins, as the presence of 2-OG is strongly synergetic towards ATP binding, while antagonistic towards ADP binding.

7. Ca^{2+} signaling in cyanobacteria

We were able to rule out that cyanobacterial PII proteins are able to sense Ca^{2+} ions, and probably the previously observed role of Ca^{2+} in acclimation of cyanobacteria to nitrogen starvation conditions is through fine-tuning the transcription of nitrogen related genes, not through the PII protein. Indeed, a recent transcriptomic study demonstrated that Ca^{2+} ions play a prominent role in the regulation of nitrogen and carbon metabolism in the cyanobacterium *Anabaena* sp. PCC 7120 to maintain the C:N balance in the cell (Walter et al. 2016). However, the PII protein was not among the regulated genes under different Ca^{2+} treatments (Walter et al. 2016), which supports our results that PII has nothing to do with Ca^{2+} sensing in cyanobacteria.
Nonetheless, in collaboration with Turku University, we were able to identify in a novel Ca^{2+} sensor EF-hand (CSE) protein in cyanobacteria. The transcriptome and biophysical approaches indicated that CSE is an important protein for regulation of photosynthesis, the correct assembly of phycobilisomes (PBS) and downstream energy and electron transfer routes, thereby affecting the overall cell fitness. The CSE protein presumably exerts its function via manipulating the free intracellular Ca^{2+} concentration depending on its cellular abundance, which in return might affect

the protein assembly of the photosynthetic complexes. A possible signaling function of CSE as well is to interact with different protein targets for posttranslational modification of PBS proteins, however this hypothesis needs to be experimentally addressed in the future.

H. Highlights and future outlooks

The PII superfamily consists of widespread signal transduction proteins found in all domains of life. In general, canonical PII proteins sense the cellular energy state through competitive binding of ATP and ADP and they sense the C/N balance through binding of 2-OG; an intermediate of the TCA cycle. Phylogenetically, PII-like proteins are clearly related to the classical PII family but lack the signature sequences and the functions for most of them are unknown. With this work, we opened the gate for investigation of PII-like proteins, in particular in cyanobacteria. The overall results of this current work put a new piece to the understanding of the crosstalk between carbon and nitrogen metabolisms in cyanobacteria with identification of cell signaling PII-like protein SbtB, which work together with the canonical PII portion for maintaining the C:N balance, where PII controls mainly the N-metabolism and SbtB controls the C-metabolism (Fig. 19). Our physiological, biochemical, and structural results revealed that SbtB is a novel trimeric cAMP sensor protein with a ferredoxin-like fold and a part of the cyanobacterial CCM system for sensing the fluctuation of inorganic carbon levels, thereby linking the cellular inorganic carbon responses with cAMP signaling. We were able to show as well that the cyanobacterial PII-like protein CutA is structurally trimeric, however it is not involved in heavy metal tolerance at least in cyanobacteria.

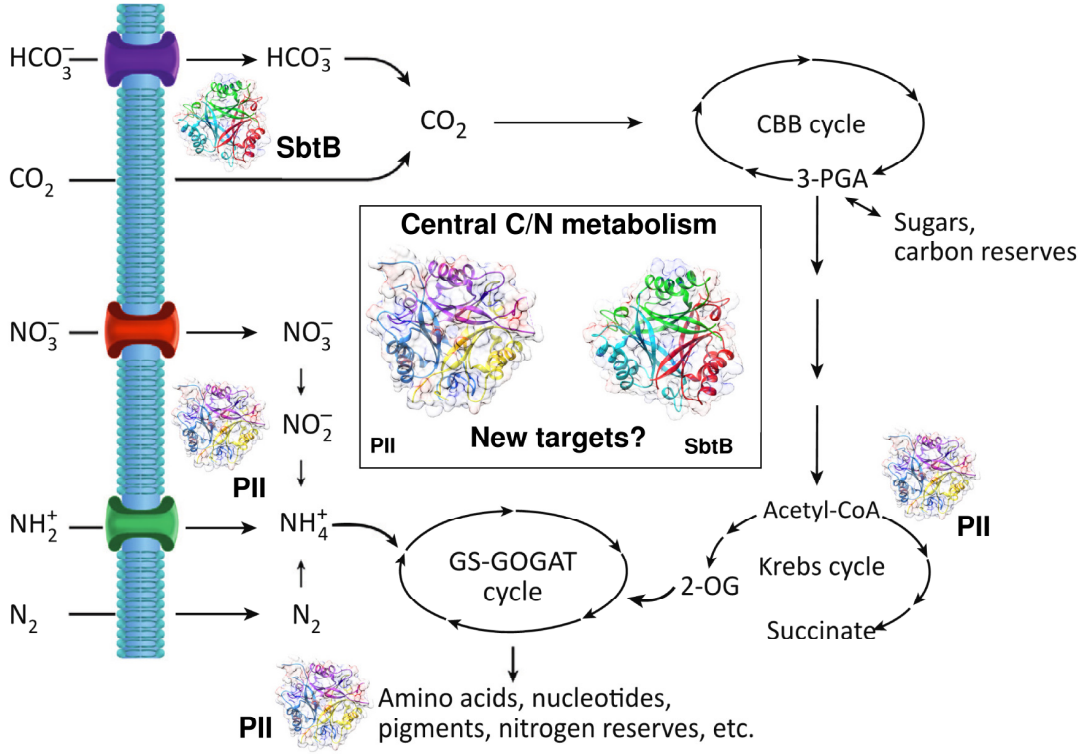


Fig. 19. Modulation of carbon and nitrogen crosstalk by cell signaling: PII and PII-like SbtB proteins, modified from (Zhang et al. 2018).

For canonical PII proteins, we were able to show a conserved role of PII in the regulation of the committed step in the arginine biosynthesis, through interacting with NAGK in Rhodophyta and Chlorophyta. In the nonphotosynthetic alga *P. parva*, due to the evolution of heterotrophic metabolism, it appears that PII and NAGK coevolved together into a stable hetero-oligomeric complex, with PII as a subunit of the enzyme. The hetero PII-NAGK complex is specialized to tune NAGK activity in a highly sensitive manner in response to arginine and glutamine levels, where PII senses glutamine and NAGK senses arginine. Additionally, the PII signaling system in Rhodophyta seems to represent an evolutionary intermediate between prokaryotic cyanobacterial and Chlorophyta PII systems. For the sensory properties of PII proteins, it seems that the higher plants lost in later stages of the evolution the ability to sense ADP and 2-OG. Apparently, it seems that the Lys58 of cyanobacterial PII is important for anchoring the T- and B-loops for efficient signaling and complexing with NAGK and PipX proteins. Moreover, we ruled out the possibility of the PII protein to sense Ca^{2+} directly in cyanobacteria, instead we identified a new conserved Ca^{2+} binding protein in cyanobacteria, which apparently plays a role in the regulation of photosynthetic activity via influencing on the assembly of phycobilisomes complexes.

Future outlooks

1. Our bioinformatic analysis revealed several DUF clusters with expected trimeric PII-fold, thus it would be interesting to characterize the DUF clusters, in particular, DUF3240 and DUF190 which are directly linked to canonical PII and SbtB clusters, respectively, for better understanding of cellular functions of PII-like proteins and the evolution of the PII superfamily.
2. Characterization of the new subgroups of SbtB cluster (in particular the SbtB-like homologues) in the near-future would contribute for better understanding of PII superfamily evolution.
3. Determination the structural bases of SbtB in complex with ADP and ATP, to reveal whether SbtB possesses ATPase activity or not.
4. Since SbtB protein is abundant in cytoplasm implying that SbtB protein has different targets. Therefore, it would be interesting to identify the new targets of SbtB protein possibly via mass spectrometry-based pulldown assay.
5. Overexpression and purification of recombinant SbtA protein to determine biochemical properties and structural bases of SbtA-B interaction.
6. Resolving the structural bases of PII-NAGK complex from *P. parva* in comparison to *C. reinhardtii* to determine the features caused the evolution of hetero-oligomeric complex.
7. Expanding the investigation of PII-NAGK systems in other nonphotosynthetic organisms, where the photosynthesis was lost, to reveal if the coevolution of such novel heteromeric PII-NAGK complex, is a common phenomenon or it is evolved independently.
8. Determination of the structural bases of the CSE protein to determine the Ca^{2+} binding mode to the protein.

I. References

1. Agostoni M, Montgomery BL. (2014) Survival strategies in the aquatic and terrestrial world: the impact of second messengers on cyanobacterial processes. Life. 4(4):745-69.
2. Archibald JM. (2009) The puzzle of plastid evolution. Curr Biol. 19(2):R81-8.
3. Arcondéguy T, Jack R, Merrick M. (2001) P(II) signal transduction proteins, pivotal players in microbial nitrogen control. Microbiol Mol Biol Rev. 65(1):80-105.
4. Arnesano F, Banci L, Benvenuti M, Bertini I, Calderone V, Mangani S, Viezzoli MS. (2003) The evolutionarily conserved trimeric structure of CutA1 proteins suggests a role in signal transduction. J Biol Chem. 278(46):45999-6006.
5. Atkinson MR, Kamberov ES, Weiss RL, Ninfa AJ. (1994) Reversible uridylylation of the *Escherichia coli* PII signal transduction protein regulates its ability to stimulate the dephosphorylation of the transcription factor nitrogen regulator I (NRI or NtrC). J Biol Chem. 269(45):28288-93.
6. Bandyopadhyay A, Arora A, Jain S, Laskar A, Mandal C, Ivanisenko VA, Fomin ES, Pintus SS, Kolchanov NA, Maiti S, Ramachandran S. (2010) Expression and molecular characterization of the *Mycobacterium tuberculosis* PII protein. J Biochem. 147(2):279-89.
7. Beez S, Fokina O, Herrmann C, Forchhammer K. (2009) N-acetyl-L-glutamate kinase (NAGK) from oxygenic phototrophs: P(II) signal transduction across domains of life reveals novel insights in NAGK control. J Mol Biol. 389(4):748-58.
8. Bekker A, Holland HD, Wang PL, Rumble D 3rd, Stein HJ, Hannah JL, Coetzee LL, Beukes NJ. (2004) Dating the rise of atmospheric oxygen. Nature. 427(6970):117-20.
9. Binder A. (1982) Respiration and photosynthesis in energy-transducing membranes of cyanobacteria. J Bioenerg Biomembr. 14(5-6):271-86.
10. Blankenship RE, Hartman H. (1998) The origin and evolution of oxygenic photosynthesis. Trends Biochem Sci. 23(3):94-7.
11. Bueno R, Pahel G, Magasanik B. (1985) Role of glnB and glnD gene products in regulation of the glnALG operon of *Escherichia coli*. J Bacteriol. 164(2):816-22.
12. Burkhead JL, Abdel-Ghany SE, Morrill JM, Pilon-Smits EA, Pilon M. (2003) The Arabidopsis thaliana CUTA gene encodes an evolutionarily conserved copper binding chloroplast protein. Plant J. 34(6):856-67.
13. Burnap RL, Hagemann M, Kaplan A. (2015) Regulation of CO_2 concentrating mechanism in cyanobacteria. Life. 5(1):348-71.
14. Burnap RL, Nambudiri R, Holland S. (2013) Regulation of the carbon-concentrating mechanism in the cyanobacterium *Synechocystis* sp. PCC6803 in response to changing light intensity and inorganic carbon availability. Photosynth Res. 118(1-2):115-24.
15. Caldovic L, Tuchman M. (2003) N-acetylglutamate and its changing role through evolution. Biochem J. 372(Pt 2):279-90.

16. Campeotto I, Zhang Y, Mladenov MG, Freemont PS, Gründling A. (2015) Complex structure and biochemical characterization of the *Staphylococcus aureus* cyclic diadenylate monophosphate (c-di-AMP)-binding protein PstA, the founding member of a new signal transduction protein family. J Biol Chem. 290(5):2888-901.
17. Cann MJ, Hammer A, Zhou J, Kanacher T. (2003) A defined subset of adenylyl cyclases is regulated by bicarbonate ion. J Biol Chem. 278(37):35033-8.
18. Castellen P, Rego FG, Portugal ME, Benelli EM. (2011) The *Streptococcus mutans* GlnR protein exhibits an increased affinity for the glnRA operon promoter when bound to GlnK. Braz J Med Biol Res. 44(12):1202-8.
19. Chellamuthu VR, Alva V, Forchhammer K. (2013) From cyanobacteria to plants: conservation of PII functions during plastid evolution. Planta. 237(2):451-62.
20. Chellamuthu VR, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, Hartmann MD, Forchhammer K. (2014) A widespread glutamine-sensing mechanism in the plant kingdom. Cell. 159(5):1188-1199.
21. Chellamuthu VR. (2013) Novel structures of PII signal transduction proteins from oxygenic phototropic organisms. PhD Dissertation, Tübingen University, Germany.
22. Chen Y, Cann MJ, Litvin TN, Iourgenko V, Sinclair ML, Levin LR, Buck J. (2000) Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. Science. 289(5479):625-8.
23. Cho Y, Sharma V, Sacchettini JC. (2003) Crystal structure of ATP phosphoribosyltransferase from *Mycobacterium tuberculosis*. J Biol Chem. 278(10):8333-9.
24. Commichau FM, Forchhammer K, Stülke J. (2006) Regulatory links between carbon and nitrogen metabolism. Curr Opin Microbiol. 9(2):167-72.
25. Conroy MJ, Durand A, Lupo D, Li XD, Bullough PA, Winkler FK, Merrick M. (2007) The crystal structure of the Escherichia coli AmtB-GlnK complex reveals how GlnK regulates the ammonia channel. Proc Natl Acad Sci U S A. 104(4):1213-8.
26. Coutts G, Thomas G, Blakey D, Merrick M. (2002) Membrane sequestration of the signal transduction protein GlnK by the ammonium transporter AmtB. EMBO J. 21(4):536-45.
27. da Rocha RA, Weschenfelder TA, de Castilhos F, de Souza EM, Huergo LF, Mitchell DA. (2013) Mathematical model of the binding of allosteric effectors to the *Escherichia coli* PII signal transduction protein GlnB. Biochemistry. 52(15):2683-93.
28. Daley SM, Kappell AD, Carrick MJ, Burnap RL. (2012) Regulation of the cyanobacterial CO_2-concentrating mechanism involves internal sensing of NADP+ and α-ketogutarate levels by transcription factor CcmR. PLoS One. 7(7):e41286.
29. Detsch C, Stülke J. (2003) Ammonium utilization in *Bacillus subtilis*: transport and regulatory functions of NrgA and NrgB. Microbiology. 149(Pt 11):3289-97.

30. Du J, Förster B, Rourke L, Howitt SM, Price GD. (2014) Characterisation of cyanobacterial bicarbonate transporters in *E. coli* shows that SbtA homologs are functional in this heterologous expression system. PLoS One. 9(12):e115905.
31. Eisenhut M, Ruth W, Haimovich M, Bauwe H, Kaplan A, Hagemann M. (2008) The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proc Natl Acad Sci U S A. 105(44):17199-204.
32. Ermilova E, Lapina T, Zalutskaya Z, Minaeva E, Fokina O, Forchhammer K. (2013) PII signal transduction protein in Chlamydomonas reinhardtii: localization and expression pattern. Protist. 164(1):49-59.
33. Espinosa J, Forchhammer K, Burillo S, Contreras A. (2006) Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate a manner with PII and NtcA. Mol Microbiol. 61(2):457-69.
34. Espinosa J, Forchhammer K, Contreras A. (2007) Role of the *Synechococcus* PCC 7942 nitrogen regulator protein PipX in NtcA-controlled processes. Microbiology. 153(Pt 3):711-8.
35. Espinosa J, Rodríguez-Mateos F, Salinas P, Lanza VF, Dixon R, de la Cruz F, Contreras A. (2014) PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. Proc Natl Acad Sci U S A. 111(23):E2423-30.
36. Fedorova K, Kayumov A, Woyda K, Ilinskaja O, Forchhammer K. (2013) Transcription factor TnrA inhibits the biosynthetic activity of glutamine synthetase in *Bacillus subtilis*. FEBS Lett. 587(9):1293-8.
37. Feria Bourrellier AB, Valot B, Guillot A, Ambard-Bretteville F, Vidal J, Hodges M. (2010) Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. Proc Natl Acad Sci U S A. 107(1):502-7.
38. Ferrario-Méry S, Meyer C, Hodges M. (2008) Chloroplast nitrite uptake is enhanced in *Arabidopsis* PII mutants. FEBS Lett. 582(7):1061-6.
39. Figge RM, Cassier-Chauvat C, Chauvat F, Cerff R. (2001) Characterization and analysis of an NAD(P)H dehydrogenase transcriptional regulator critical for the survival of cyanobacteria facing inorganic carbon starvation and osmotic stress. Mol Microbiol. 39(2):455-68.
40. Fokina O, Chellamuthu VR, Forchhammer K, Zeth K. (2010a) Mechanism of 2-oxoglutarate signaling by the Synechococcus elongatus PII signal transduction protein. Proc Natl Acad Sci U S A. 107(46):19760-5.
41. Fokina O, Chellamuthu VR, Zeth K, Forchhammer K. (2010b) A novel signal transduction protein P(II) variant from *Synechococcus elongatus* PCC 7942 indicates a two-step process for NAGK-P(II) complex formation. J Mol Biol. 399(3):410-21.
42. Fokina O, Herrmann C, Forchhammer K. (2011) Signal-transduction protein P(II) from *Synechococcus elongatus* PCC 7942 senses low adenylate energy charge *in vitro*. Biochem J. 440(1):147-56.
43. Fong ST, Camakaris J, Lee BT. (1995) Molecular genetics of a chromosomal locus involved in copper tolerance in *Escherichia coli* K-12. Mol Microbiol. 15(6):1127-37.

44. Forcada-Nadal A, Llácer JL, Contreras A, Marco-Marín C, Rubio V. (2018) The PII-NAGK-PipX-NtcA Regulatory Axis of Cyanobacteria: A Tale of Changing Partners, Allosteric Effectors and Non-covalent Interactions. Front Mol Biosci. 5:91.
45. Forcada-Nadal A, Palomino-Schätzlein M, Neira JL, Pineda-Lucena A, Rubio V. (2017). The PipX Protein, When not bound to its targets, has its signaling C-terminal helix in a flexed conformation. Biochemistry. 56(25):3211-3224.
46. Forchhammer K, Hedler A, Strobel H, Weiss V. (1999) Heterotrimerization of PII-like signalling proteins: implications for PII-mediated signal transduction systems. Mol Microbiol. 33(2):338-49.
47. Forchhammer K, Lüddecke J. (2016) Sensory properties of the PII signalling protein family. FEBS J. 283(3):425-37.
48. Forchhammer K, Tandeau de Marsac N. (1994) The PII protein in the cyanobacterium *Synechococcus* sp. strain PCC 7942 is modified by serine phosphorylation and signals the cellular N-status. J Bacteriol. 176(1):84-91.
49. Forchhammer K. (2004) Global carbon/nitrogen control by PII signal transduction in cyanobacteria: from signals to targets. FEMS Microbiol Rev. 28(3):319-33.
50. Forchhammer K. (2008) P(II) signal transducers: novel functional and structural insights. Trends Microbiol. 16(2):65-72.
51. Forchhammer K. (2010) The network of P(II) signalling protein interactions in unicellular cyanobacteria. Adv Exp Med Biol. 675:71-90.
52. Fujishiro T, Ermler U, Shima S. (2014) A possible iron delivery function of the dinuclear iron center of HcgD in [Fe]-hydrogenase cofactor biosynthesis. FEBS Lett. 588(17):2789-93.
53. Gancedo JM. (2013) Biological roles of cAMP: variations on a theme in the different kingdoms of life. Biol Rev Camb Philos Soc. 88(3):645-68.
54. Gerhardt EC, Araújo LM, Ribeiro RR, Chubatsu LS, Scarduelli M, Rodrigues TE, Monteiro RA, Pedrosa FO, Souza EM, Huergo LF. (2012) Influence of the ADP/ATP ratio, 2-oxoglutarate and divalent ions on *Azospirillum brasilense* PII protein signalling. Microbiology. 158(6):1656-63.
55. Gerhardt EC, Rodrigues TE, Müller-Santos M, Pedrosa FO, Souza EM, Forchhammer K, Huergo LF. (2015) The bacterial signal transduction protein GlnB regulates the committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of acetyl-CoA carboxylase. Mol Microbiol. 95(6):1025-35.
56. Godsey MH, Minasov G, Shuvalova L, Brunzelle JS, Vorontsov II, Collart FR, Anderson WF. (2007) The 2.2 A resolution crystal structure of *Bacillus cereus* Nif3-family protein YqfO reveals a conserved dimetal-binding motif and a regulatory domain. Protein Sci. 16(7):1285-93.
57. Gould SB, Waller RF, McFadden GI. (2008) Plastid evolution. Annu Rev Plant Biol. 59:491-517.
58. Gundlach J, Dickmanns A, Schröder-Tittmann K, Neumann P, Kaesler J, Kampf J, Herzberg C, Hammer E, Schwede F, Kaever V, Tittmann K, Stülke J, Ficner R. (2015) Identification,

characterization, and structure analysis of the cyclic di-AMP-binding PII-like signal transduction protein DarA. J Biol Chem. 290(5):3069-80.

59. Hackenberg C, Huege J, Engelhardt A, Wittink F, Laue M, Matthijs HC, Kopka J, Bauwe H, Hagemann M. (2012) Low-carbon acclimation in carboxysome-less and photorespiratory mutants of the cyanobacterium *Synechocystis* sp. strain PCC 6803. Microbiology. 158(Pt 2):398-413.
60. Hagemann M, Eisenhut M, Hackenberg C, Bauwe H. (2010) Pathway and importance of photorespiratory 2-phosphoglycolate metabolism in cyanobacteria. Adv Exp Med Biol. 675:91-108.
61. Hammer A, Hodgson DR, Cann MJ. (2006) Regulation of prokaryotic adenylyl cyclases by CO_2. Biochem J. 396(2):215-8.
62. Hauf K, Kayumov A, Gloge F, Forchhammer K. (2016a) The Molecular basis of TnrA control by glutamine synthetase in *Bacillus subtilis*. J Biol Chem. 291(7):3483-95.
63. Hauf W, Schmid K, Gerhardt EC, Huergo LF, Forchhammer K. (2016b) Interaction of the nitrogen regulatory protein GlnB (PII) with biotin carboxyl carrier protein (BCCP) controls acetyl-CoA levels in the cyanobacterium *Synechocystis* sp. PCC 6803. Front Microbiol. 7:1700.
64. Hayes JM. (1996) The earliest memories of life on Earth. Nature. 384(6604):21-2.
65. Heinrich A, Maheswaran M, Ruppert U, Forchhammer K. (2004) The *Synechococcus elongatus* P signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. Mol Microbiol. 52(5):1303-14.
66. Heinrich A, Woyda K, Brauburger K, Meiss G, Detsch C, Stülke J, Forchhammer K. (2006) Interaction of the membrane-bound GlnK-AmtB complex with the master regulator of nitrogen metabolism TnrA in Bacillus subtilis. J Biol Chem. 281(46):34909-17.
67. Helfmann S, Lü W, Litz C, Andrade SL. (2010) Cooperative binding of MgATP and MgADP in the trimeric P(II) protein GlnK2 from *Archaeoglobus fulgidus*. J Mol Biol. 402(1):165-77.
68. Hennon GMM, Ashworth J, Groussman RD, Berthiaume C, Morales RL, Baliga NS, Orellana MV, Armbrust EV. (2015) Diatom acclimation to elevated CO_2 via cAMP signaling and coordinated gene expression. Nat Clim Chang, 5:761–765.
69. Hesketh A, Fink D, Gust B, Rexer HU, Scheel B, Chater K, Wohlleben W, Engels A. (2002) The GlnD and GlnK homologues of *Streptomyces coelicolor* A3(2) are functionally dissimilar to their nitrogen regulatory system counterparts from enteric bacteria. Mol Microbiol. 46(2):319-30.
70. Hohmann-Marriott MF, Blankenship RE. (2011) Evolution of photosynthesis. Annu Rev Plant Biol. 62:515-48.
71. Holland HD. (2006) The oxygenation of the atmosphere and oceans. Philos Trans R Soc Lond B Biol Sci. 361(1470):903-15.
72. Hsieh MH, Lam HM, van de Loo FJ, Coruzzi G. (1998) A PII-like protein in Arabidopsis: putative role in nitrogen sensing. Proc Natl Acad Sci U S A. 95(23):13965-70.
73. Huergo LF, Chandra G, Merrick M. (2013) P(II) signal transduction proteins: nitrogen regulation and beyond. FEMS Microbiol Rev. 37(2):251-83.

74. Huergo LF, Chubatsu LS, Souza EM, Pedrosa FO, Steffens MB, Merrick M. (2006a) Interactions between PII proteins and the nitrogenase regulatory enzymes DraT and DraG in *Azospirillum brasilense*. FEBS Lett. 580(22):5232-6.
75. Huergo LF, Dixon R. (2015) The emergence of 2-Oxoglutarate as a master regulator metabolite. Microbiol Mol Biol Rev. 79(4):419-35.
76. Huergo LF, Merrick M, Monteiro RA, Chubatsu LS, Steffens MB, Pedrosa FO, Souza EM. (2009) In vitro interactions between the PII proteins and the nitrogenase regulatory enzymes dinitrogenase reductase ADP-ribosyltransferase (DraT) and dinitrogenase reductase-activating glycohydrolase (DraG) in Azospirillum brasilense. J Biol Chem. 284(11):6674-82.
77. Huergo LF, Merrick M, Pedrosa FO, Chubatsu LS, Araujo LM, Souza EM. (2007) Ternary complex formation between AmtB, GlnZ and the nitrogenase regulatory enzyme DraG reveals a novel facet of nitrogen regulation in bacteria. Mol Microbiol. 66(6):1523-35.
78. Huergo LF, Souza EM, Araujo MS, Pedrosa FO, Chubatsu LS, Steffens MB, Merrick M. (2006b) ADP-ribosylation of dinitrogenase reductase in *Azospirillum brasilense* is regulated by AmtB-dependent membrane sequestration of DraG. Mol Microbiol. 59(1):326-37.
79. Irmler A, Forchhammer K. (2001) A PP2C-type phosphatase dephosphorylates the PII signaling protein in the cyanobacterium *Synechocystis* PCC 6803. Proc Natl Acad Sci U S A. 98(23):12978-83.
80. Irmler A, Sanner S, Dierks H, Forchhammer K. (1997) Dephosphorylation of the phosphoprotein P(II) in *Synechococcus* PCC 7942: identification of an ATP and 2-oxoglutarate-regulated phosphatase activity. Mol Microbiol. 26(1):81-90.
81. Jiang P, Mayo AE, Ninfa AJ. (2007) Escherichia coli glutamine synthetase adenylyltransferase (ATase, EC 2.7.7.49): kinetic characterization of regulation by PII, PII-UMP, glutamine, and alpha-ketoglutarate. Biochemistry. 46(13):4133-46.
82. Jiang P, Ninfa AJ. (2007) *Escherichia coli* PII signal transduction protein controlling nitrogen assimilation acts as a sensor of adenylate energy charge in vitro. Biochemistry. 46(45):12979-96.
83. Jiang P, Ninfa AJ. (2009) α-Ketoglutarate Controls the Ability of the *Escherichia coli* PII Signal Transduction Protein To Regulate the Activities of NRII (NtrB) but Does Not Control the Binding of PII to NRII. Biochemistry. 48(48): 11514–11521.
84. Jiang P, Peliska JA, Ninfa AJ. (1998) Enzymological characterization of the signal-transducing uridylyltransferase/uridylyl-removing enzyme (EC 2.7.7.59) of Escherichia coli and its interaction with the PII protein. Biochemistry. 37(37):12782-94.
85. Jiang P, Ventura AC, Ninfa AJ. (2012) Characterization of the reconstituted UTase/UR-PII-NRII-NRI bicyclic signal transduction system that controls the transcription of nitrogen-regulated (Ntr) genes in *Escherichia coli*. Biochemistry. 51(45):9045-57.
86. Jiang YL, Wang XP, Sun H, Han SJ, Li WF, Cui N, Lin GM, Zhang JY, Cheng W, Cao DD, Zhang ZY, Zhang CC, Chen Y, Zhou CZ. (2018) Coordinating carbon and nitrogen metabolic signaling through the cyanobacterial global repressor NdhR. Proc Natl Acad Sci U S A. 115(2):403-408.

87. Karl D, Letelier R, Tupas L, Dore J, Christian J, Hebel D. (1997) The role of nitrogen fixation in biogeochemical cycling in the subtropical North Pacific Ocean. Nature. **388**:533–538
88. Kayumov A, Heinrich A, Fedorova K, Ilinskaya O, Forchhammer K. (2011) Interaction of the general transcription factor TnrA with the PII-like protein GlnK and glutamine synthetase in *Bacillus subtilis*. FEBS J. 278(10):1779-89.
89. Kayumov A, Heinrich A, Sharipova M, Iljinskaya O, Forchhammer K. (2008) Inactivation of the general transcription factor TnrA in Bacillus subtilis by proteolysis. Microbiology. 154:2348-55.
90. Kerr RA. (2005) Earth science. The story of O_2. Science. 308(5729):1730-2.
91. Khademi S, O'Connell J 3rd, Remis J, Robles-Colmenares Y, Miercke LJ, Stroud RM. (2004) Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 A. Science. 305(5690):1587-94.
92. Klähn S, Orf I, Schwarz D, Matthiessen JK, Kopka J, Hess WR, Hagemann M. (2015) Integrated transcriptomic and metabolomic characterization of the low-carbon response using an ndhR mutant of *Synechocystis* sp. PCC 6803. Plant Physiol. 169(3):1540-56.
93. Kloft N, Forchhammer K. (2005) Signal transduction protein PII phosphatase PphA is required for light-dependent control of nitrate utilization in *Synechocystis* sp. strain PCC 6803. J Bacteriol. 187(19):6683-90.
94. Kloft N, Rasch G, Forchhammer K. (2005) Protein phosphatase PphA from *Synechocystis* sp. PCC 6803: the physiological framework of PII-P dephosphorylation. Microbiology. 151(Pt 4):1275-83.
95. Labella JI, Obrebska A, Espinosa J, Salinas P, Forcada-Nadal A, Tremiño L, Rubio V, Contreras A. (2016) Expanding the Cyanobacterial Nitrogen Regulatory Network: The GntR-Like Regulator PlmA Interacts with the PII-PipX Complex. Front Microbiol. 7:1677.
96. Lee HM, Flores E, Forchhammer K, Herrero A, Tandeau De Marsac N. (2000) Phosphorylation of the signal transducer PII protein and an additional effector are required for the PII-mediated regulation of nitrate and nitrite uptake in the Cyanobacterium *Synechococcus* sp. PCC 7942. Eur J Biochem. 267(2):591-600.
97. Leganés F, Forchhammer K, Fernández-Piñas F. (2009) Role of calcium in acclimation of the cyanobacterium *Synechococcus elongatus* PCC 7942 to nitrogen starvation. Microbiology. 155(Pt 1):25-34.
98. Liang D, Nunes-Tavares N, Xie HQ, Carvalho S, Bon S, Massoulié J. (2009) Protein CutA undergoes an unusual transfer into the secretory pathway and affects the folding, oligomerization, and secretion of acetylcholinesterase. J Biol Chem. 284(8):5195-207.
99. Llácer JL, Contreras A, Forchhammer K, Marco-Marín C, Gil-Ortiz F, Maldonado R, Fita I, Rubio V. (2007) The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. Proc Natl Acad Sci U S A. 104(45):17644-9.

100. Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K, Rubio V. (2010) Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. Proc Natl Acad Sci U S A. 107(35):15397-402.

101. Lohkamp B, McDermott G, Campbell SA, Coggins JR, Lapthorn AJ. (2004) The structure of Escherichia coli ATP-phosphoribosyltransferase: identification of substrate binding sites and mode of AMP inhibition. J Mol Biol. 336(1):131-44.

102. Long BM, Badger MR, Whitney SM, Price GD. (2007) Analysis of carboxysomes from *Synechococcus* PCC7942 reveals multiple Rubisco complexes with carboxysomal proteins CcmM and CcaA. J Biol Chem. 282(40):29323-35.

103. Long BM, Rae BD, Rolland V, Förster B, Price GD. (2016) Cyanobacterial CO_2-concentrating mechanism components: function and prospects for plant metabolic engineering. Curr Opin Plant Biol. 31:1-8.

104. Lüddecke J, Forchhammer K. (2015) Energy Sensing versus 2-Oxoglutarate Dependent ATPase Switch in the Control of *Synechococcus* PII Interaction with Its Targets NAGK and PipX. PLoS One. 10(8):e0137114.

105. Lüddecke J. (2016) Development of FRET Based Biosensors for *in vitro* and *in vivo* Metabolite Quantification. PhD Dissertation, Tübingen University, Germany.

106. Luque I, Forchhammer K. (2008) Nitrogen assimilation and C/N balance sensing. Chapter 13 in The Cyanobacteria: Molecular Biology, Genomics and Evolution (Book edited by: Herrero A, Flores E), Caister Academic Press, UK.

107. Lyons TW, Reinhard CT, Planavsky NJ. (2014) The rise of oxygen in Earth's early ocean and atmosphere. Nature. 506(7488):307-15.

108. Ma CW, Lüddecke J, Forchhammer K, Zeng AP. (2014) Population shift of binding pocket size and dynamic correlation analysis shed new light on the anticooperative mechanism of PII protein. Proteins. 82(6):1048-59.

109. Ma W, Ogawa T. (2015) Oxygenic photosynthesis-specific subunits of cyanobacterial NADPH dehydrogenases. IUBMB Life. 67(1):3-8.

110. Maier S, Schleberger P, Lü W, Wacker T, Pflüger T, Litz C, Andrade SL. (2011) Mechanism of disruption of the Amt-GlnK complex by P(II)-mediated sensing of 2-oxoglutarate. PLoS One. 6(10):e26327.

111. Merrick M. (2015) Post-translational modification of PII signal transduction proteins. Front Microbiol. 5:763.

112. Minaeva E, Forchhammer K, Ermilova E. (2015) Glutamine assimilation and feedback regulation of L-acetyl-N-glutamate kinase activity in *Chlorella variabilis* NC64A results in changes in arginine pools. Protist. 166(5):493-505.

113. Mizuno Y, Moorhead GB, Ng KK. (2007) Structural basis for the regulation of N-acetylglutamate kinase by PII in *Arabidopsis thaliana*. J Biol Chem. 282(49):35733-40.

114. Mojzsis SJ, Arrhenius G, McKeegan KD, Harrison TM, Nutman AP, Friend CR. (1996) Evidence for life on Earth before 3,800 million years ago. Nature. 384(6604):55-9.

115. Molina J, Hazzouri KM, Nickrent D, Geisler M, Meyer RS, Pentony MM, Flowers JM, Pelser P, Barcelona J, Inovejas SA, Uy I, Yuan W, Wilkins O, Michel CI, Locklear S, Concepcion GP, Purugganan MD. (2014) Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). Mol Biol Evol. 31(4):793-803.

116. Müller M, Hopfner KP, Witte G. (2015) c-di-AMP recognition by *Staphylococcus aureus* PstA. FEBS Lett. 589(1):45-51.

117. Muro-Pastor MI, Reyes JC, Florencio FJ. (2001) Cyanobacteria perceive nitrogen status by sensing intracellular 2-oxoglutarate levels. J Biol Chem. 276(41):38320-8.

118. Muro-Pastor MI, Reyes JC, Florencio FJ. (2005) Ammonium assimilation in cyanobacteria. Photosynth Res. 83(2):135-50.

119. Ninfa AJ, Jiang P. (2005) PII signal transduction proteins: sensors of alpha-ketoglutarate that regulate nitrogen metabolism. Curr Opin Microbiol. 8(2):168-73.

120. Nishimura T, Takahashi Y, Yamaguchi O, Suzuki H, Maeda S, Omata T. (2008) Mechanism of low CO_2-induced activation of the cmp bicarbonate transporter operon by a LysR family protein in the cyanobacterium *Synechococcus elongatus* strain PCC 7942. Mol Microbiol. 68(1):98-109.

121. Ohashi Y, Shi W, Takatani N, Aichi M, Maeda S, Watanabe S, Yoshikawa H, Omata T. (2011) Regulation of nitrate assimilation in cyanobacteria. J Exp Bot. 62(4):1411-24.

122. Oliveira MA, Gerhardt EC, Huergo LF, Souza EM, Pedrosa FO, Chubatsu LS. (2015) 2-Oxoglutarate levels control adenosine nucleotide binding by *Herbaspirillum seropedicae* PII proteins. FEBS J. 282(24):4797-809.

123. Omata T, Price GD, Badger MR, Okamura M, Gohta S, Ogawa T. (1999) Identification of an ATP-binding cassette transporter involved in bicarbonate uptake in the cyanobacterium *Synechococcus* sp. strain PCC 7942. Proc Natl Acad Sci U S A. 96(23):13571-6.

124. Orf I, Klähn S, Schwarz D, Frank M, Hess WR, Hagemann M, Kopka J. (2015) Integrated Analysis of Engineered Carbon Limitation in a Quadruple CO_2/HCO_3^- Uptake Mutant of *Synechocystis* sp. PCC 6803. Plant Physiol. 169(3):1787-806.

125. Orf I, Schwarz D, Kaplan A, Kopka J, Hess WR, Hagemann M, Klähn S. (2016) CyAbrB2 contributes to the transcriptional regulation of low CO_2 acclimation in *Synechocystis* sp. PCC 6803. Plant Cell Physiol. 57(10):2232-2243.

126. Osanai T, Sato S, Tabata S, Tanaka K. (2005) Identification of PamA as a PII-binding membrane protein important in nitrogen-related and sugar-catabolic gene expression in *Synechocystis* sp. PCC 6803. J Biol Chem. 280(41):34684-90.

127. Palanca C, Pedro-Roig L, Llácer JL, Camacho M, Bonete MJ, Rubio V. (2014) The structure of a PII signaling protein from a halophilic archaeon reveals novel traits and high-salt adaptations. FEBS J. 281(15):3299-314.

128. Perrier AL, Cousin X, Boschetti N, Haas R, Chatel JM, Bon S, Roberts WL, Pickett SR, Massoulié J, Rosenberry TL, Krejci E. (2000) Two distinct proteins are associated with tetrameric acetylcholinesterase on the cell surface. J Biol Chem. 275(44):34260-5.

129. Price GD, Badger MR. (1991) Evidence for the role of carboxysomes in the cyanobacterial co2-concentrating mechanism. Can J Bot. 69(5):963-973.

130. Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L. (2004) Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. Proc Natl Acad Sci U S A. 101(52):18228-33.

131. Radchenko M, Merrick M. (2011) The role of effector molecules in signal transduction by PII proteins. Biochem Soc Trans. 39(1):189-94.

132. Radchenko MV, Thornton J, Merrick M. (2010) Control of AmtB-GlnK complex formation by intracellular levels of ATP, ADP, and 2-oxoglutarate. J Biol Chem. 2010 Oct 1;285(40):31037-45.

133. Radchenko MV, Thornton J, Merrick M. (2013) P(II) signal transduction proteins are ATPases whose activity is regulated by 2-oxoglutarate. Proc Natl Acad Sci U S A. 110(32):12948-53.

134. Radchenko MV, Thornton J, Merrick M. (2014) Association and dissociation of the GlnK-AmtB complex in response to cellular nitrogen status can occur in the absence of GlnK post-translational modification. Front Microbiol. 5:731.

135. Rae BD, Long BM, Badger MR, Price GD. (2013) Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. Microbiol Mol Biol Rev. 77(3):357-79.

136. Rajendran C, Gerhardt EC, Bjelic S, Gasperina A, Scarduelli M, Pedrosa FO, Chubatsu LS, Merrick M, Souza EM, Winkler FK, Huergo LF, Li XD. (2011) Crystal structure of the GlnZ-DraG complex reveals a different form of PII-target interaction. Proc Natl Acad Sci U S A. 108(47):18972-6.

137. Raven JA. (2006) Sensing inorganic carbon: CO_2 and HCO_3^-. Biochem J. 396(2):e5-7.

138. Rodionova IA, Goodacre N, Babu M, Emili A, Uetz P, Saier MH Jr. (2018) The Nitrogen Regulatory PII Protein (GlnB) and *N*-Acetylglucosamine 6-Phosphate Epimerase (NanE) Allosterically Activate Glucosamine 6-Phosphate Deaminase (NagB) in *Escherichia coli*. J Bacteriol. 200(5), e00691-17.

139. Rodrigues CM, Takita MA, Coletta-Filho HD, Olivato JC, Caserta R, Machado MA, de Souza AA. (2008) Copper resistance of biofilm cells of the plant pathogen *Xylella fastidiosa*. Appl Microbiol Biotechnol. 77(5):1145-57.

140. Rodrigues TE, Gerhardt EC, Oliveira MA, Chubatsu LS, Pedrosa FO, Souza EM, Souza GA, Müller-Santos M, Huergo LF. (2014) Search for novel targets of the PII signal transduction protein in Bacteria identifies the BCCP component of acetyl-CoA carboxylase as a PII binding partner. Mol Microbiol. 91(4):751-61.

141. Rodrigues TE, Sassaki GL, Valdameri G, Pedrosa FO, Souza EM, Huergo LF. (2018) Fatty acid biosynthesis is enhanced in *Escherichia coli* strains with deletion in genes encoding the PII signaling proteins. Arch Microbiol. 201(2):209-214.

142. Ruppert U, Irmler A, Kloft N, Forchhammer K. (2002) The novel protein phosphatase PphA from *Synechocystis* PCC 6803 controls dephosphorylation of the signalling protein PII. Mol Microbiol. 44(3):855-64.

143. Saikatendu KS, Zhang X, Kinch L, Leybourne M, Grishin NV, Zhang H. (2006) Structure of a conserved hypothetical protein SA1388 from *S. aureus* reveals a capped hexameric toroid with two PII domain lids and a dinuclear metal center. BMC Struct Biol. 6:27.

144. Sant'Anna FH, Trentini DB, de Souto Weber S, Cecagno R, da Silva SC, Schrank IS. (2009) The PII superfamily revised: a novel group and evolutionary insights. J Mol Evol. 68(4):322-36.

145. Schumacher MA, Chinnam NB, Cuthbert B, Tonthat NK, Whitfill T. (2015) Structures of regulatory machinery reveal novel molecular mechanisms controlling *B. subtilis* nitrogen homeostasis. Genes Dev. 29(4):451-64.

146. Schwarz D, Nodop A, Hüge J, Purfürst S, Forchhammer K, Michel KP, Bauwe H, Kopka J, Hagemann M. (2011) Metabolic and transcriptomic phenotyping of inorganic carbon acclimation in the Cyanobacterium *Synechococcus elongatus* PCC 7942. Plant Physiol. 155(4):1640-55.

147. Shapiro BM. (1969) The glutamine synthetase deadenylylating enzyme system from *Escherichia coli*. Resolution into two components, specific nucleotide stimulation, and cofactor requirements. Biochemistry. 8(2):659-70.

148. Shibata M, Katoh H, Sonoda M, Ohkawa H, Shimoyama M, Fukuzawa H, Kaplan A, Ogawa T. (2002) Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: function and phylogenetic analysis. J Biol Chem. 277(21):18658-64.

149. Shibata M, Ohkawa H, Kaneko T, Fukuzawa H, Tabata S, Kaplan A, Ogawa T. (2001) Distinct constitutive and low-CO_2-induced CO_2 uptake systems in cyanobacteria: genes involved and their phylogenetic relationship with homologous genes in other organisms. Proc Natl Acad Sci U S A. 98(20):11789-94.

150. Smith CS, Morrice NA, Moorhead GB.(2004) Lack of evidence for phosphorylation of *Arabidopsis thaliana* PII: implications for plastid carbon and nitrogen signaling. Biochim Biophys Acta. 1699(1-2):145-54.

151. Smith CS, Weljie AM, Moorhead GB. (2003) Molecular properties of the putative nitrogen sensor PII from *Arabidopsis thaliana*. Plant J. 33(2):353-60.

152. Smith DR, Asmail SR. (2014) Next-generation sequencing data suggest that certain nonphotosynthetic green plants have lost their plastid genomes. New Phytol. 204(1):7-11.

153. Smith DR, Hua J, Archibald JM, Lee RW. (2013) Palindromic genes in the linear mitochondrial genome of the nonphotosynthetic green alga *Polytomella magna*. Genome Biol Evol. 5(9):1661-7.

154. Smith DR, Lee RW. (2014) A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*. Plant Physiol. 164(4):1812-9.

155. Spät P, Maček B, Forchhammer K. (2015) Phosphoproteome of the cyanobacterium *Synechocystis* sp. PCC 6803 and its dynamics during nitrogen starvation. Front Microbiol. 6:248.

156. Spät P. (2017) Establishment of quantitative proteomics in cyanobacteria reveals O-phosphorylation as fundamental process for stress-acclimation to nitrogen deficiency. PhD Dissertation, Tübingen University, Germany.

157. Srivastava R, Pisareva T, Norling B. (2005) Proteomic studies of the thylakoid membrane of *Synechocystis* sp. PCC 6803. Proteomics. 5(18):4905-16.

158. Steegborn C, Litvin TN, Levin LR, Buck J, Wu H. (2005) Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment. Nat Struct Mol Biol. 12(1):32-7.

159. Strösser J, Lüdke A, Schaffer S, Krämer R, Burkovski A. (2004) Regulation of GlnK activity: modification, membrane sequestration and proteolysis as regulatory principles in the network of nitrogen control in *Corynebacterium glutamicum*. Mol Microbiol. 54(1):132-47.

160. Su J, Forchhammer K. (2013) Determinants for substrate specificity of the bacterial PP2C protein phosphatase tPphA from *Thermosynechococcus elongatus*. FEBS J. 280(2):694-707.

161. Su J, Schlicker C, Forchhammer K. (2011) A third metal is required for catalytic activity of the signal-transducing protein phosphatase M tPphA. J Biol Chem. 286(15):13481-8.

162. Takahashi Y, Yamaguchi O, Omata T. (2004) Roles of CmpR, a LysR family transcriptional regulator, in acclimation of the cyanobacterium *Synechococcus* sp. strain PCC 7942 to low-CO_2 and high-light conditions. Mol Microbiol. 52(3):837-45.

163. Tirichine L, Bowler C. (2011) Decoding algal genomes: tracing back the history of photosynthetic life on Earth. Plant J. 66(1):45-57.

164. Truan D, Bjelić S, Li XD, Winkler FK. (2014) Structure and thermodynamics of effector molecule binding to the nitrogen signal transduction PII protein GlnZ from *Azospirillum brasilense*. J Mol Biol. 426(15):2783-99.

165. Truan D, Huergo LF, Chubatsu LS, Merrick M, Li XD, Winkler FK. (2010) A new P(II) protein structure identifies the 2-oxoglutarate binding site. J Mol Biol. 400(3):531-9.

166. Tsai Y, Sawaya MR, Cannon GC, Cai F, Williams EB, Heinhorst S, Kerfeld CA, Yeates TO. (2007) Structural analysis of CsoS1A and the protein shell of the *Halothiobacillus neapolitanus* carboxysome. PLoS Biol. 5(6):e144.

167. Uhrig RG, Ng KK, Moorhead GB. (2009) PII in higher plants: a modern role for an ancient protein. Trends Plant Sci. 14(9):505-11.

168. Walter J, Lynch F, Battchikova N, Aro EM, Gollan PJ. (2016) Calcium impacts carbon and nitrogen balance in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. J Exp Bot. 67(13):3997-4008.

169. Wang HL, Postier BL, Burnap RL. (2004) Alterations in global patterns of gene expression in *Synechocystis* sp. PCC 6803 in response to inorganic carbon limitation and the inactivation of ndhR, a LysR family regulator. J Biol Chem. 279(7):5739-51.

170. Watzer B, Engelbrecht A, Hauf W, Stahl M, Maldener I, Forchhammer K. (2015) Metabolic pathway engineering using the central signal processor PII. *Microb Cell Fact.* 14:192.

171. Wheatley NM, Eden KD, Ngo J, Rosinski JS, Sawaya MR, Cascio D, Collazo M, Hoveida H, Hubbell WL, Yeates TO. (2016) A PII-Like protein regulated by bicarbonate: structural and biochemical studies of the carboxysome-associated CPII protein. J Mol Biol. 428(20):4013-4030.

172. Woodger FJ, Badger MR, Price GD. (2005) Sensing of inorganic carbon limitation in *Synechococcus* PCC7942 is correlated with the size of the internal inorganic carbon pool and involves oxygen. Plant Physiol. 139(4):1959-69.

173. Xu Y, Carr PD, Clancy P, Garcia-Dominguez M, Forchhammer K, Florencio F, Vasudevan SG, Tandeau de Marsac N, Ollis DL. (2003) The structures of the PII proteins from the cyanobacteria *Synechococcus* sp. PCC 7942 and *Synechocystis* sp. PCC 6803. Acta Crystallogr D Biol Crystallogr. 59(Pt 12):2183-90.

174. Yang J, Li Q, Yang H, Yan L, Yang L, Yu L. (2008) Overexpression of human CUTA isoform2 enhances the cytotoxicity of copper to HeLa cells. Acta Biochim Pol. 55(2):411-5.

175. Zehr JP, Waterbury JB, Turner PJ, Montoya JP, Omoregie E, Steward GF, Hansen A, Karl DM. (2001) Unicellular cyanobacteria fix N_2 in the subtropical North Pacific Ocean. Nature. 412(6847):635-8.

176. Zeth K, Fokina O, Forchhammer K. (2014) Structural basis and target-specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. J Biol Chem. 289(13):8960-72.

177. Zhang CC, Zhou CZ, Burnap RL, Peng L. (2018) Carbon/Nitrogen Metabolic Balance: Lessons from Cyanobacteria. Trends Plant Sci. 23(12):1116-1130.

178. Zhang Y, Pohlmann EL, Serate J, Conrad MC, Roberts GP. (2010) Mutagenesis and functional characterization of the four domains of GlnD, a bifunctional nitrogen sensor protein. J Bacteriol. 192(11):2711-21.

179. Zhang Y, Pu H, Wang Q, Cheng S, Zhao W, Zhang Y, Zhao J. (2007) PII is important in regulation of nitrogen metabolism but not required for heterocyst formation in the Cyanobacterium *Anabaena* sp. PCC 7120. J Biol Chem. 282(46):33641-8.

180. Zhao MX, Jiang YL, Xu BY, Chen Y, Zhang CC, Zhou CZ. (2010) Crystal structure of the cyanobacterial signal transduction protein PII in complex with PipX. J Mol Biol. 402(3):552-9.

181. Zhao Y, Shi Y, Zhao W, Huang X, Wang D, Brown N, Brand J, Zhao J. (2005) CcbP, a calcium-binding protein from *Anabaena* sp. PCC 7120, provides evidence that calcium ions regulate heterocyst differentiation. Proc Natl Acad Sci U S A. 102(16):5744-8.

J. Appendex

1. Accepted publications

1.1. Publication 1 (Research Article):

Lapina T[#], **Selim KA**[#], Forchhammer K, Ermilova E. (2018) The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins. Sci Rep. 8(1):790.

1.2. Publication 2 (Research Article):

Selim KA, Haase F, Hartmann MD, Hagemann M, Forchhammer K. (2018) PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. Proc Natl Acad Sci U S A. 115(21): E4861-E4869.

1.3. Publication 3 (Research Article):

Selim KA, Lapina T, Forchhammer K, Ermilova E. (2019) Interaction of N-acetyl-l-glutamate kinase with the PII signal transducer in the non-photosynthetic alga *Polytomella parva*: Co-evolution towards a hetero-oligomeric enzyme. FEBS J. doi: 10.1111/febs.14989.

1.4. Publication 4 (Research Article):

Walter J, **Selim KA**, Leganés F, Fernández-Piñas F, Vothknecht UC, Forchhammer K, Aro E-M, Gollan PJ. (2019) A novel Ca^{2+}-binding protein influences photosynthetic electron transport in *Anabaena* sp. PCC 7120. Biochim Biophys Acta Bioenerg. 1860(6):519-532.

1.5. Publication 5 (Research Article):

Selim KA, Haffner M, Watzer B, Forchhammer K. (2019) Tuning the *in vitro* sensing and signaling properties of cyanobacterial PII protein by mutation of key residues. Sci Rep. 9(1): 18985.

2. Submitted manuscript

2.1. Manuscript 1 (Research Article):

Selim KA, Alva V, Hartmann MD, Forchhammer K. Structural and functional characterization of cyanobacterial PII-like protein CutA does not hint at an involvement in heavy metal tolerance. FEBS J. (Submitted 2019).

SCIENTIFIC REPORTS

OPEN

The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins

Received: 8 November 2017
Accepted: 15 December 2017
Published online: 15 January 2018

Tatyana Lapina[1], Khaled A. Selim[2], Karl Forchhammer[2] & Elena Ermilova[1]

PII superfamily consists of widespread signal transduction proteins found in all domains of life. Whereas they are well-studied in Archaea, Bacteria and Chloroplastida, no PII homolog has been analyzed in Rhodophyta (red algae), where PII is encoded by a chloroplast localized *glnB* gene. Here, we characterized relevant sensory properties of PII from the red alga *Porphyra purpurea* (PpPII) in comparison to PII proteins from different phyla of oxygenic phototrophs (cyanobacteria, *Chlamydomonas* and *Physcomitrella*) to assess evolutionary conservation versus adaptive properties. Like its cyanobacterial counterparts, PpPII binds ATP/ADP and 2-oxoglutarate in synergy with ATP. However, green algae and land plant PII proteins lost the ability to bind ADP. In contrast to PII proteins from green algae and land plants, PpPII enhances the activity of N-acetyl-L-glutamate kinase (NAGK) and relieves it from feedback inhibition by arginine in a glutamine-independent manner. Like PII from Chloroplastida, PpPII is not able to interact with the cyanobacterial transcriptional co-activator PipX. These data emphasize the conserved role of NAGK as a major PII-interactor throughout the evolution of oxygenic phototrophs, and confirms the specific role of PipX for cyanobacteria. Our results highlight the PII signaling system in red algae as an evolutionary intermediate between Cyanobacteria and Chlorophyta.

The PII superfamily were originally described as widely distributed members of a family of cell signaling proteins occurring in all domains of life[1–3] with representatives in almost all bacteria and in nitrogen-fixing archaea[4,5] as well as in oxygenic eukaryotic phototrophs[6]. The canonical PII proteins are the master regulator of nitrogen metabolism and they are encoded by *glnB* and *glnK* genes[2,7]. The superfamily of PII-like proteins was enlarged by including members that are characterized by the typical structural architecture of PII proteins but lack the typical PROSITE signature pattern of initially characterized PII proteins[7]. Those PII homologues, which contain the typically conserved PROSITE motifs of GlnB/GlnK-like PII proteins are referred as canonical PII proteins.

The canonical PII proteins have fundamentals roles as energy/carbon/nitrogen sensors[8]. The binding of small effector molecules to PII proteins allows modulation of different cellular functions. Competitive binding of ATP or ADP and synergistic binding of 2-oxoglutarate (2-OG) with ATP enables PII to estimate the current energy and nitrogen/carbon status of the cells. The various effector molecule binding events cause signaling through conformational changes within the PII tirmer, which in turn allows PII to bind to different interacting partners to regulate the actual metabolic situation. Under conditions of high 2-OG levels (poor nitrogen supply), the ATP-dependent binding of 2-OG to PII causes strong conformational changes in the T-loop, which in turn impairs the interaction of PII proteins with different targets[7]. In all examined cases studied so far, PII proteins coordinate the central C/N anabolic metabolism by regulating the activity of transcription factors, key metabolic enzymes and transporters[9–12].

In many bacteria, PII proteins of the GlnB subfamily are engaged in glutamine synthetase control at various levels[13,14] whereas PII proteins of the GlnK subfamily regulate ammonia transporter Amt by reversibly clogging

[1]Biological Faculty, Saint-Petersburg State University, Universitetskaya nab. 7/9, Saint-Petersburg, 199034, Russia. [2]Interfaculty Institute of Microbiology and Infection Medicine/Organismic Interactions Department, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 28, 72076, Tübingen, Germany. Tatyana Lapina and Khaled A. Selim contributed equally to this work. Correspondence and requests for materials should be addressed to K.F. (email: karl.forchhammer@uni-tuebingen.de) or E.E. (email: e.ermilova@spbu.ru)

its NH_3 channel[11]. In cyanobacteria, PII proteins are present in all known species. In contrast to PII proteins from other bacterial lineages, cyanobacterial PII proteins have evolved to regulate N-acetyl-L-glutamate kinase (EC 2.7.2.8) (NAGK), which is the controlling enzyme of arginine biosynthesis[15], and the general nitrogen control transcription factor NtcA through binding to the transcriptional co-activator PipX[16–18]. In common with other bacterial PII proteins, cyanobacterial PII proteins interact with the BCCP subunit of the acetyl-CoA carboxylase enzyme complex[19] to control cellular acetyl-CoA levels.

Eukaryotic homologues of canonical PII proteins are restricted to members of the plant kingdom. In Chloroplastida (green algae and land plants) they are nuclear-encoded and, in Rhodophyta they are coded by the plastid genome[6]. Genomic information from red algae revealed that PII signaling has been lost in some families of the red algae[3,6] whereas it is present in three members of Bangiophyceae family, *Porphyra purpurea, Porphyra umbilicalis* and *Pyropia yezoensis*[20,21]. It is of note that the chloroplasts of red algae have more archaic properties than those of the Chloroplastida and are probably more similar to the cyanobacterial endosymbiont that gave rise to chloroplasts. Many genes that were transferred from the chloroplast to the nucleus during the evolution of Chlorophyta are still encoded in the red algae chloroplast, such as the *glnB* gene (encoding PII) or the *argB* gene (encoding NAGK) and can thus be considered as evolutionary relics. In Chloroplastida, PII and NAGK are localized in the chloroplast[22,23], where the PII proteins control the ornithine pathway via activity regulation of NAGK, as in cyanobacteria[24]. However, in green algae and land plants, NAGK activity-regulation responds - via a specific feature of the respective PII proteins - to the cellular glutamine levels in addition to the primary effects ATP and 2-OG[24,25]. A short additional C-terminal segment, absent in bacterial PII proteins, was shown to act as low-affinity glutamine binding site. This segment seems to have evolved when the *glnB* gene from the endosymbiont was translocated to the nucleus, since it is absent in Rhodophyta, where this translocation has not taken place. This suggests that the PII signaling protein from red algae resembles the PII signaling system of the early endosymbiont. However, functional analysis of red algal PII proteins is necessary to close the gap in our understanding of PII evolution in oxygenic phototrophs.

P. purpurea has been one of the important reference red algae for genomic and biochemical studies of many cellular processes including nitrogen metabolism[20,26–28]. For this reason, we started studying PII signaling in *P. purpurea*, in particular with respect to the sensory binding of small effector molecules, and compared it to the binding properties of PII proteins across different phyla representing the evolution of oxygenic phototrophs, from the cyanobacterium *Synechococcus elongatus* PCC 7942 (SyPII), through green algal *Chlamydomonas reinhardtii* (CrPII) to the moss plant *Physcomitrella patens* (PhyscoPII). Furthermore, we report PpPII-mediated NAGK regulation, with the question in mind, whether it is resembles NAGK control in cyanobacteria or green plants. Special attention was given to examine the possibility that PpPII might still be able to control the cyanobacterial PipX transcription co-activator, which is absent from Chlorophyta.

Results

Sequence based alignments of PII proteins. The predicted full-length PpPII polypeptide (P51254) encoded by the *P. purpurea GlnB* gene has the typical length bacterial PII proteins (112 amino acids), corresponding to a molecular weight of 12.320 kDa. A sequence alignment with PII proteins from other red algae, bacteria and Chloroplastida is shown in Fig. 1. The highest degree of identity occurs with the cyanobacterial homologues *S. elongatus* PCC 7942 (63.96%) and *Synechocystis* sp. PCC 6803 (62.26%). The PpPII identity with plant PII proteins was lower and demonstrated 45.37% (*C. reinhardtii*), 52.29% (*Arabidopsis thaliana* and *Oryza sativa Japonica*) and 51.38% (*P. patens* and *Solanum lycopersicum*). Similar to PII homologs of bacteria, PpPII does not contain the N-terminal transit signal peptides and the unique C-terminal segment that are present in proteins of Chloroplastida[24] (Fig. 1).

Regions of extremely high local identities are known as PII signature patterns that have been defined at PROSITE PS00496 and PS00638 (Fig. 1). The signature pattern I (Y-[KR]-G-[AS]-[AE]-Y) contains the residues that are involved in the interaction with NAGK. The second pattern which was found in all prokaryotic PII-proteins comprises the B-loop region [ST]-x(3)-G-[DY]-G-[KR]-[IV]-[FW]-[LIVM], and is also conserved in PpPII. This motif is involved in the binding of the phosphates of adenyl nucleotides in the effector molecule binding cleft of the PII trimers. Examination of the aligned sequences shows that residues necessary for effector molecule binding are conserved. However, since the detailed binding characteristics cannot be predicted from the mere presence of conserved residues, the effector-binding properties should be analysed.

Binding of effectors molecular to PpPII protein. To gain additional insights into the ligand binding properties of PpPII in comparison to cyanobacterial and chloroplastid PII proteins, isothermal titration calorimetry (ITC) experiments were performed using highly purified Strep-tagged or His-tagged recombinant PpPII, SyPII, CrPII, and PhyscoPII proteins (Supplementary Fig. S1) with different protein/ligand/ concentrations to determine the best fitting conditions of the binding model. The titration experiments were repeated with different protein preparations to confirm the reproducibility of the results. The raw isothermal data were fitted into one-binding site as well as three-sequential binding sites to define the best fitting model. Fitting using a one binding site model states an average K_d value for all available binding sites and the mean of the stoichiometry of bound ligands. However, it was reported before that the optimal fitting for different members of PII superfamily was obtained only when using a sequential binding sites model with defining three consecutives binding sites[9,29].

Table 1 and Fig. 2 show the dissociation constants and the binding isotherms for binding of the small effector molecules to red algal PpPII protein. Under optimal binding conditions, robust binding of ATP and ADP was detectable. As shown in Fig. 2A and B, PpPII protein exhibited high affinity toward ATP and ADP in μM range. When the data are fitted according to a model assuming independent binding sites, the average K_d value for all binding sites of PpPII protein bound to ATP or ADP were about 22.0 and 7.2 μM, respectively, and with a respective stoichiometry of 2.5 and 2.7 per PII trimer, respectively. Data fitting using three sequential binding sites

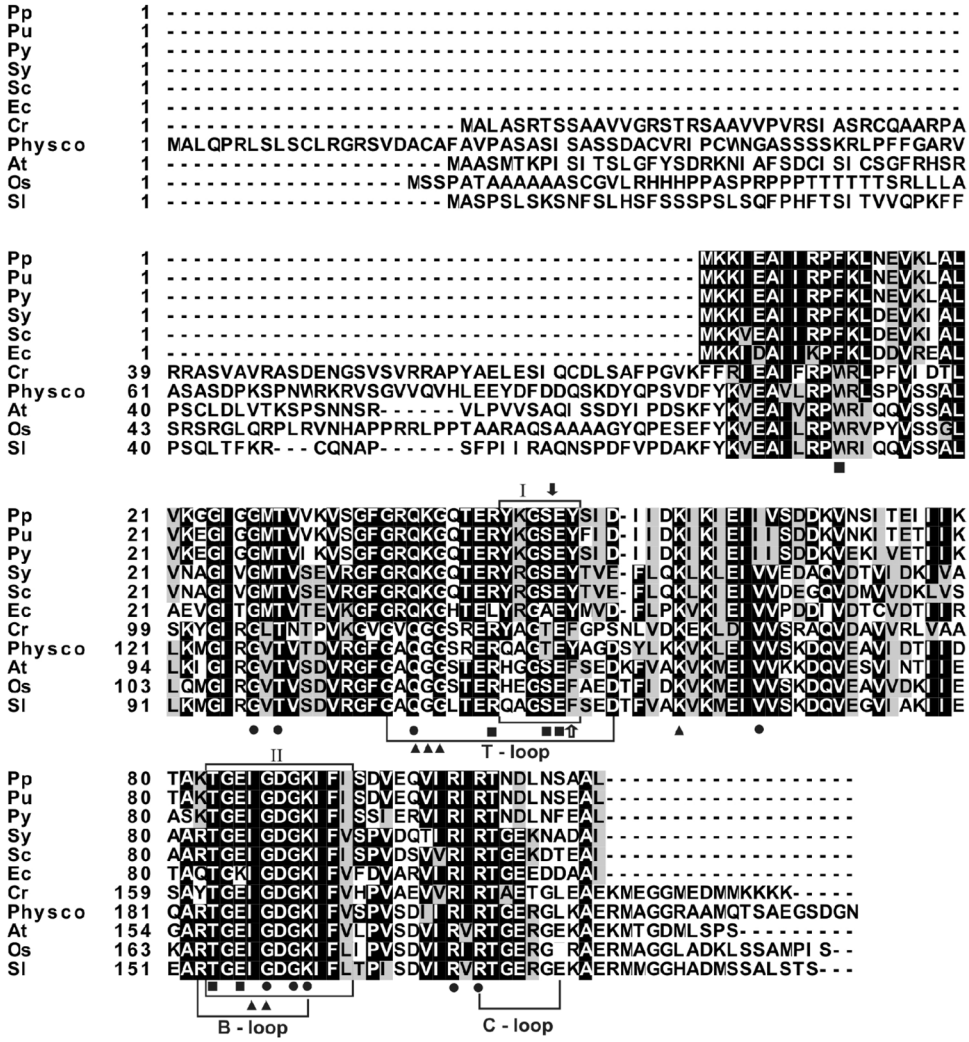


Figure 1. Comparison of the deduced amino acid sequences of PII polypeptides from red algae, bacteria, and Chloroplastida (green algae and higher plants). Aligned are deduced PII protein sequences from red algae: *Porphyra purpurea* (Pp; NP_053864.1), *Porphyra umbilicalis* (Pu; AFC39923.1), and *Pyropia yezoensis* (Py; AGH27579.1), bacteria: *Synechococcus elongatus* PCC 7942 (Sy; P0A3F4.1), *Synechocystis* sp. PCC 6803 (Sc; CAA66127.1), and *E. coli* (Ec; CAQ32926.1.), green algae: *Chlamydomonas reinhardtii* (Cr; A8JI83), and higher plants: *Physcomitrella patens* (Physco; BAF36548.1), *Arabidopsis thaliana* (At; NP_192099.1), *Oryza sativa Japonica* (Os, NP_001054562.1), and *Solanum lycopersicum* (Sl, AAR14689.1). Residues highlighted in black are identical or conserved in at least 55% of all aligned PII proteins. Amino acids in a gray background represent similar residues. Box I and box II refer to PII signature patterns I and II, respectively. The positions of the Tyr residue that is uridylylated in *E. coli* PII and Ser residue that is phosphorylated in *Synechoccocus* PCC 7942 are indicated by white and solid black arrows, respectively. ATP-(●), NAGK- (■) and 2-OG-binding residues (▲) are highlighted. Alignments were made with the ClustalW program and refined manually.

model revealed two high-affinity sites (sites one and two) and a third low-affinity site. This analysis resolved a low K_{d1} value for the first binding site of 0.5 µM for ATP, and 3.4 µM for ADP. For binding site 2, the K_{d2} values were 18.7 µM for ATP and 15.2 µM for ADP, and for binding sites 3, K_{d3} of 141.6 µM for ATP and 131.8 µM for ADP (Table 1). The binding affinities towards ATP and ADP are very similar, reflecting that the ADP binds to the algal PpPII protein almost as efficient as ATP.

Both, binding of ATP and ADP, display a strong anticooperativity among the three consecutive sites. Compared to the adenyl nucleotide binding properties of the previously analysed cyanobacterial SyPII protein[9], the overall characteristics are similar. However, in the PpPII protein, the extent of anticooperativity is even higher, spanning a range of more than two orders of magnitude, whereas in SyPII, the range is only one order of magnitude (from 4 µM to 47 µM for ATP and from 10 to 133 µM for ADP)[9]. For binding site one, there is a clear preference for ATP, whereas in the other two sites, competition between ATP and ADP will occur.

For 2-OG binding to PpPII protein, strong binding isotherm was noticeable in the presence of ATP (Fig. 2E). Optimal fitting was obtained assuming three sequential binding sites model, again revealing strong anti-cooperativity. The first site was occupied at low 2-OG concentrations of K_{d1} 2.7 µM, the second site requires

Titrant/Protein	One-site binding model		Three-sites binding model		
	Average K_d (μM)	N*	K_{d1} (μM)	K_{d2} (μM)	K_{d3} (μM)
Red algal PpPII					
ATP	22.0 ± 10.8	2.5	0.5 ± 0.2	18.7 ± 9.2	141.6 ± 80.8
ADP	7.2 ± 2.5	2.7	3.4 ± 0.5	15.2 ± 0.7	131.8 ± 2.7
ATP (in presence 2-OG)	0.2 ± 0.3	1.0	1.4 ± 1.8	1.0 ± 0.9	83.9 ± 62.4
ADP (in presence 2-OG)	10.8 ± 2.8	2.7	10.3 ± 10.0	43.6 ± 4.7	125.9 ± 0.9
2-OG (in presence ATP)	4.0 ± 2.4	1.0	2.7 ± 1.4	50.7 ± 16.1	80.0 ± 22.0
2-OG (in presence ADP)	No binding				
Cyanobacterial SyPII					
ATP	43.3 ± 37.7	3.2	7.5 ± 7.3	15.9 ± 12.2	85.4 ± 74.6
ADP[9]	NF	NF	(10.6 ± 3.2)	(19.3 ± 2.3)	(133.4 ± 5.2)
ATP (in presence 2-OG)	2.7 ± 0.1	1.8	0.5 ± 0.5	3.0 ± 1.7	325.5 ± 171.0
ADP (in presence 2-OG and 150 μM ADP)	45.8 ± 44.0	4.0	10.7 ± 1.9	26.4 ± 4.2	100.9 ± 102.5
2-OG (in presence ATP)	21.1 ± 6.8	1.4	2.9 ± 4.0	6.4 ± 2.7	83.0 ± 43.7
2-OG (in presence ADP)[49]	(No binding)				

Table 1. Dissociation constants (K_d) for ATP, ADP and 2-OG as indicated binding to recombinant red algal PpPII and cyanobacterial SyPII proteins, respectively. The raw data were fitted using one-site and three sequential binding sites models for PII trimer. The K_d values correspond to the mean of the independent experiments ± SD. For comparison, data for ADP and 2-OG in presence of ADP binding to SyPII are given in parentheses. *N, number of calculated binding sites; NF, not fitted.

approximately 19-fold higher concentrations of 2-OG, corresponding to a K_{d2} of 50.7 μM. Occupation of the third binding site took place with a K_{d3} of 80.0 μM. These values are quite similar to the properties of cyanobacterial SyPII[9]. No binding of 2-OG could be detected in presence of ADP, in contrast to what was reported for plant PII protein from *A. thaliana*[30].

Next, we investigated the influence of 2-OG on the binding of ATP or ADP to PpPII. As shown for several bacterial PII proteins, the presence of 2-OG highly increased the affinity towards ATP binding (Table 1 and Fig. 2C). By contrast, in the presence of 2-OG the binding enthalpy for ADP was strongly decreased when compared with the titration in the absence of 2-OG (Fig. 2, compare D with B). Therefore, the presence of 2-OG increased the K_d values of the first and the second ADP binding sites (Table 1).

To study comparatively the sensing properties of PII proteins in the course of the endosymbiont (chloroplast) evolution, we choose the PII protein from the cyanobacterium *S. elongatus* (SyPII) (Table 1), from the green alga *C. reinhardtii* (CrPII) and from the higher eukaryotic moss-plant *P. patens* (PhyscoPII) (Table 2). As reported previously, SyPII protein was able to bind ATP, ADP and 2-OG, with 2-OG being bound exclusively in the presence of ATP with negative cooperativity for the three-available binding sites[9,31]. As a control, we obtained comparable results with the previously reported K_d values for ATP and 2-OG binding in presence of ATP (Table 1, Fig. 3A and B)[9]. Moreover, in the presence of 2-OG the binding enthalpy for ATP was strongly increased as compared to binding of ATP alone (Fig. 3, compare C with A). Thus, like in other PII proteins, the presence of 2-OG enhances the affinity for ATP (Table 1). Furthermore, the effect of 2-OG against ADP on SyPII binding was investigated. First, ADP was titrated to SyPII protein in the presence of 1 mM 2-OG. The titration curves showed a complex biphasic curve; in the first injections, the enthalpy increased, until a maximum was reached at ADP concentration of 80.0 ± 7.1 μM (Fig. 3D). Subsequently, the heat change signals decreased again, indicating gradual saturation of PII by ADP (Fig. 3D). Therefore, it seems that elevated 2-OG levels antagonistically prevent formation of the PII:ADP complex. To test whether a certain threshold level of ADP can relieve the inhibitory effect of 2-OG on ADP binding to SyPII (suggested by the biphasic curve), ADP was titrated again against SyPII protein in presence of 1 mM 2-OG and 150 μM ADP (Fig. 3E). A typical curve was obtained again which could be fitted with sequential binding (Table 1).

Table 2 and Fig. 4 show the dissociation constants and the binding isotherms for the binding of the small effector molecules to green algal CrPII protein. Interestingly, we were unable to notice any binding towards ATP and ADP (Fig. 4A and B). However, the crystal structure of CrPII in complex with ATP and 2-OG was solved previously (PDB: 4USI)[24]. Therefore, we tested ATP binding in the presence of 2-OG, which revealed a strong ATP binding with high affinity in μM range (Fig. 4C). This suggests that the binding affinities toward ATP under the tested conditions are very low and the ATP affinities increase dramatically in presence of 2-OG. As expected, no ADP binding to CrPII protein was detected in presence of 2-OG (Fig. 4D). Conversely, we observe binding of 2-OG in presence of ATP but never observed binding of 2-OG in presence of ADP (Fig. 4E and F).

Finally, we assayed the plant PII protein from moss *P. patens* (PhyscoPII) for binding of ATP, ADP and 2-OG. We were able to detect ATP binding (Table 2, Fig. 5A). Data fitting using three sequential binding sites model revealed high-affinity for the first binding site (K_d of 5.7 μM) and anticooperativeity on the second and third sites with K_d values of 13.6 and 48.2 μM, respectively. Surprisingly, we were unable to detect any binding for ADP, nor for 2-OG in the presence of ATP (Table 2, Fig. 5B,D and E). Consistent with this, the presence of 2-OG did not change the affinity for ATP (Table 2, and Fig. 5C). Thus, unlike the other PII proteins, 2-OG seems to have no effect on the PhyscoPII protein.

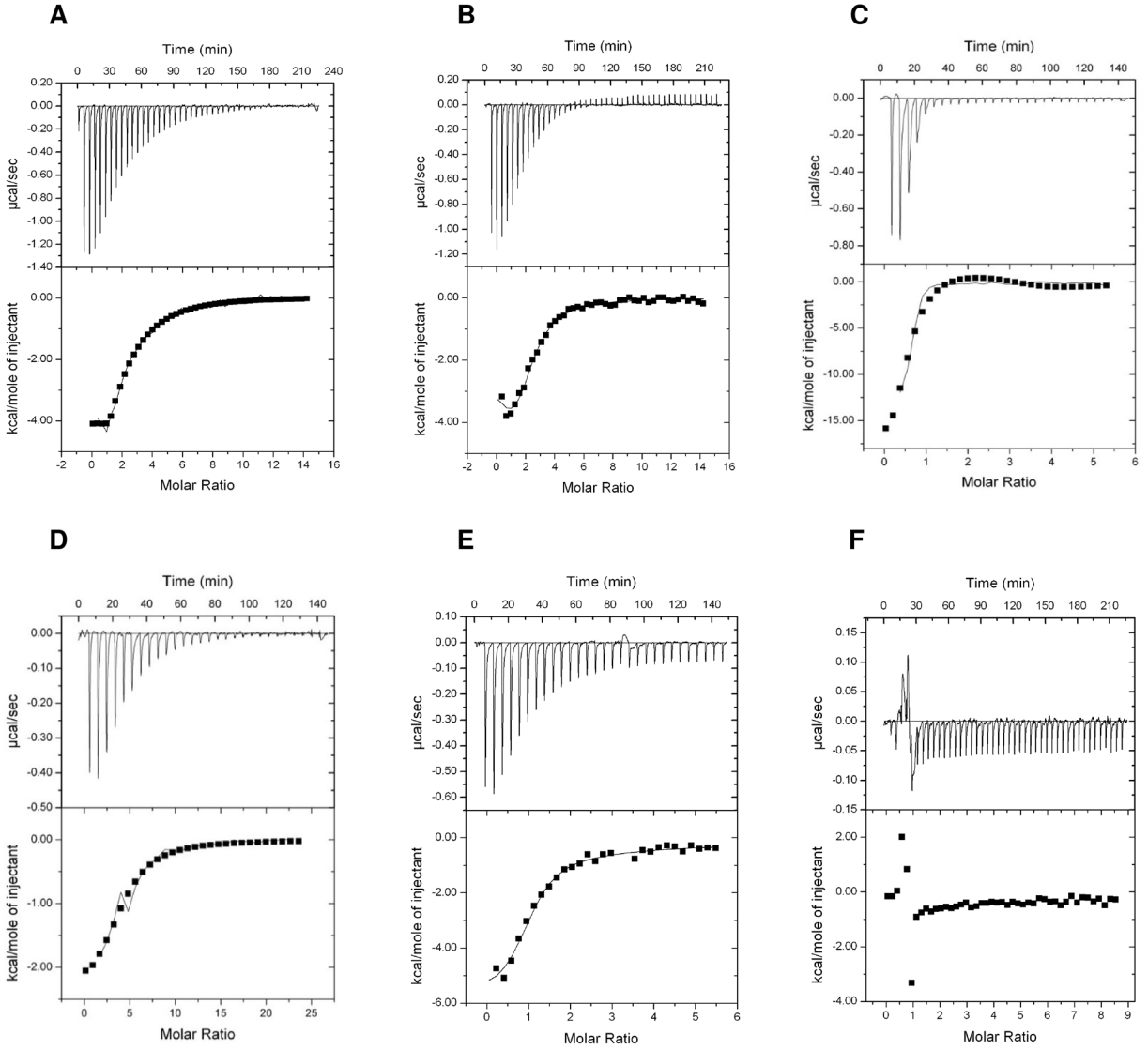


Figure 2. ITC analysis of binding small effector molecules to red algal PpPII protein. The upper panels show the raw data in the form of the heat effect during the titration of PpPII solution (trimer concentration) with ligands. The lower panels show the binding isotherm and the best-fit curve according to the three sequential binding sites model. Titration of PpPII protein (**A**) In 28.3 µM with 2 mM ATP. (**B**) In 28.3 µM with 2 mM ADP. (**C**) In 20 µM with 0.5 mM ATP in presence of 1 mM 2-OG. (**D**) In 9 µM with 1 mM ADP in presence of 1 mM 2-OG. (**E**) In 23.3 µM with 1 mM 2-OG in the presence of 1 mM ATP. (**F**) In 23.3 µM with 1 mM 2-OG in the presence of 1 mM ADP.

Enzymatic properties of the NAGK-PpPII complex. PpPII shared the signature residues involved in NAGK-PII interaction (Fig. 1). The experiments shown above (Table 1, Fig. 2) indicated that the PpPII protein responded to the effector molecules ATP, ADP and 2-OG quite similarly to cyanobacterial PII. Since, the NAGK is highly conserved enzyme across different domains of life. The amino acid sequences of NAGK enzymes are sharing high identity across different phyla (Supplementary Fig. S2). Earlier, we showed that the signal transduction PII proteins across different domains of life able to regulate NAGK activity of different phyla[32]. Therefore, we wanted subsequently to investigate the putative interaction between the red algal PpPII and cyanobacterial NAGK, and the ability of PpPII protein to activate NAGK. A recombinant NAGK from *Synechocystis* sp. PCC 6803 (ScNAGK) was expressed with its N-terminus fused to a His_6 tag. The kinetic constants of the purified recombinant ScNAGK enzyme showed a K_m value for N-acetyl glutamate (NAG) of 16.23 ± 2.14 mM and a K_{cat} of $8.93 \pm 1.40\ s^{-1}$. In the presence of PpPII (Fig. 6), the apparent K_m for NAG dropped to 4.25 ± 0.36, whereas the K_{cat} remained almost unchanged with a value to $10.02 \pm 1.11\ s^{-1}$. This corresponds to a catalytic efficiency (K_{cat}/K_m) of 0.55×10^3 and $2.36 \times 10^3\ s^{-1}M^{-1}$ for free and PpPII-complexed ScNAGK, respectively. The data showed that complex formation with PpPII caused an about 4-fold increase of the overall NAGK catalytic efficiency.

Titrant/Protein	One-site binding model		Three-sites binding model		
	Average K_d (μM)	N*	K_{d1} (μM)	K_{d2} (μM)	K_{d3} (μM)
Green algal CrPII					
ATP	No binding				
ADP	No binding				
ATP (in presence 2-OG)	35.8 ± 1.6	2.6	5.8 ± 5.8	53.1 ± 58.3	61.6 ± 9.7
ADP (in presence 2-OG)	No binding				
2-OG (in presence ATP)	15.0 ± 5.4	2.6	4.2 ± 0.7	21.3 ± 11.0	45.5 ± 19.8
2-OG (in presence ADP)	No binding				
Plant-moss PhyscoPII					
ATP	26.8 ± 10.5	2.9	5.7 ± 1.4	13.6 ± 7.2	48.2 ± 29.3
ADP	No binding				
ATP (in presence 2-OG)	17.6 ± 7.5	5.3	8.0 ± 6.0	16.3 ± 11.4	128.8 ± 128.1
ADP (in presence 2-OG)	No binding				
2-OG (in presence ATP)	No binding				

Table 2. Dissociation constants (K_d) for ATP, ADP and 2-OG as indicated binding to recombinant green algal CrPII and plant-moss PhyscoPII proteins, respectively. The raw data were fitted using one-site and three sequential binding sites models for PII trimer. The K_d values correspond to the mean of the independent experiments ± SD. *N, number of calculated binding sites.

Biochemical characterization of interaction between PpPII and ScNAGK. The relief from arginine inhibition by PII-NAGK complex formation is the major checkpoint for metabolic control of the arginine biosynthesis pathway in cyanobacteria and Chloroplastida[15,24]. To determine whether PpPII affected the catalytic activity of ScNAGK, we assayed NAGK by a coupled assay[32]. Feedback inhibition by arginine occurred with a half maximal inhibitory concentration (IC_{50}) of 6.2 μM (Fig. 7A). Addition of PpPII protein to ScNAGK raised the IC_{50} for arginine by approximately 4-fold. As expected, the presence of PpPII changed arginine inhibition of ScNAGK without the presence of glutamine.

Like other canonical PII proteins, PpPII sensed 2-OG (Fig. 2). In cyanobacteria and plants, binding of 2-OG usually antagonizes the interaction of PII with NAGK. To determine next the response of PpPII-ScNAGK interaction towards 2-OG, an assay was set up, where PpPII and ScNAGK (at a ratio of 5 PII trimers to one NAGK hexamer) were incubated together with arginine and titrated with increasing concentrations of 2-OG. In these assays, the mixture contained 0.01 mM arginine that shows relatively low inhibition for the PpPII-ScNAGK complex but already processes highly inhibitory for free ScNAGK (Fig. 7A). As shown in Fig. 7B, addition of 2-OG to the PpPII–ScNAGK–arginine mixture indeed inhibited ScNAGK activity in a concentration-dependent manner. Half-maximal inhibition (IC_{50}) of NAGK activity was attained at 0.22 ± 0.01 mM 2-OG. This concentration additionally reflected the affinity of PpPII towards the effector molecule 2-OG and indicates that all three binding sites have to be occupied by 2-OG in order to abrogate the productive interaction with NAGK.

Competition between ATP and ADP at different ratios affects PpPII-mediated activation of NAGK. All available structural information of PII proteins revelas that ATP and ADP bind to the same binding sites[2,7]. When ATP and ADP are simultaneously present, they will compete for binding to the PII trimer[31]. Since, the ATP/ADP binding site of the red algal PpPII protein is highly conserved (Fig. 1), competition between the nucleotides will occur. Since the activation of NAGK by PII is dependent on the ATP-ligated state of PII[9] and ADP inhibits PII-NAGK complex formation[9,31], competition between ATP/ADP for PII binding can be monitored by measuring NAGK activity[31] under different ATP/ADP ratios. For this assay, we used an AGPR-coupled activity assay for NAGK[33], which allows measurement of NAGK activity in presence of ADP. Different concentrations of ADP were added to reaction mixtures containing either 1 mM or 2 mM ATP, in the absence or presence of PpPII (Fig. 8). Addition of ADP led to a monotonical and shallow decrease in the activity of ScNAGK at both fixed ATP concentration. In the presence of PpPII, the effect by the addition of ADP was more pronounced (Fig. 8B). A steep decrease was observed up to 2 mM ADP, indicating that complex formation with NAGK is sensitive towards elevated ADP levels.

At the highest ADP/ATP ratio (4 mM ADP with 1 mM ATP), ScNAGK activity in the presence of PpPII was as low as ScNAGK activity in the absence of PpPII (Fig. 8). These responses were very similar to those of *S. elongatus* NAGK-PII complex to variable ATP/ADP levels[31].

Investigation of the PipX-PpPII interaction. In cyanobacteria, PipX is the second known receptor of PII signaling[18]. We questioned whether red algal PII proteins have lost or retained the ability to bind to PipX. To test this possibility, we first analyzed complex formation using pull-down experiments (Supplementary Fig. S3A and B). The PII protein from *Synechocystis* sp. PCC 6803 (ScPII) was mixed with PipX in the presence of ADP and then the Strep-tagged PII was immobilized on the Strep-Tactin II column (Supplementary Fig. S3A). After extensive washes, the proteins were eluted with desthiobiotin and analyzed by Tricine-SDS-PAGE. Both proteins were detected in the elution. Thus, confirming previous studies[16–18], cyanobacterial PII was bound to PipX.

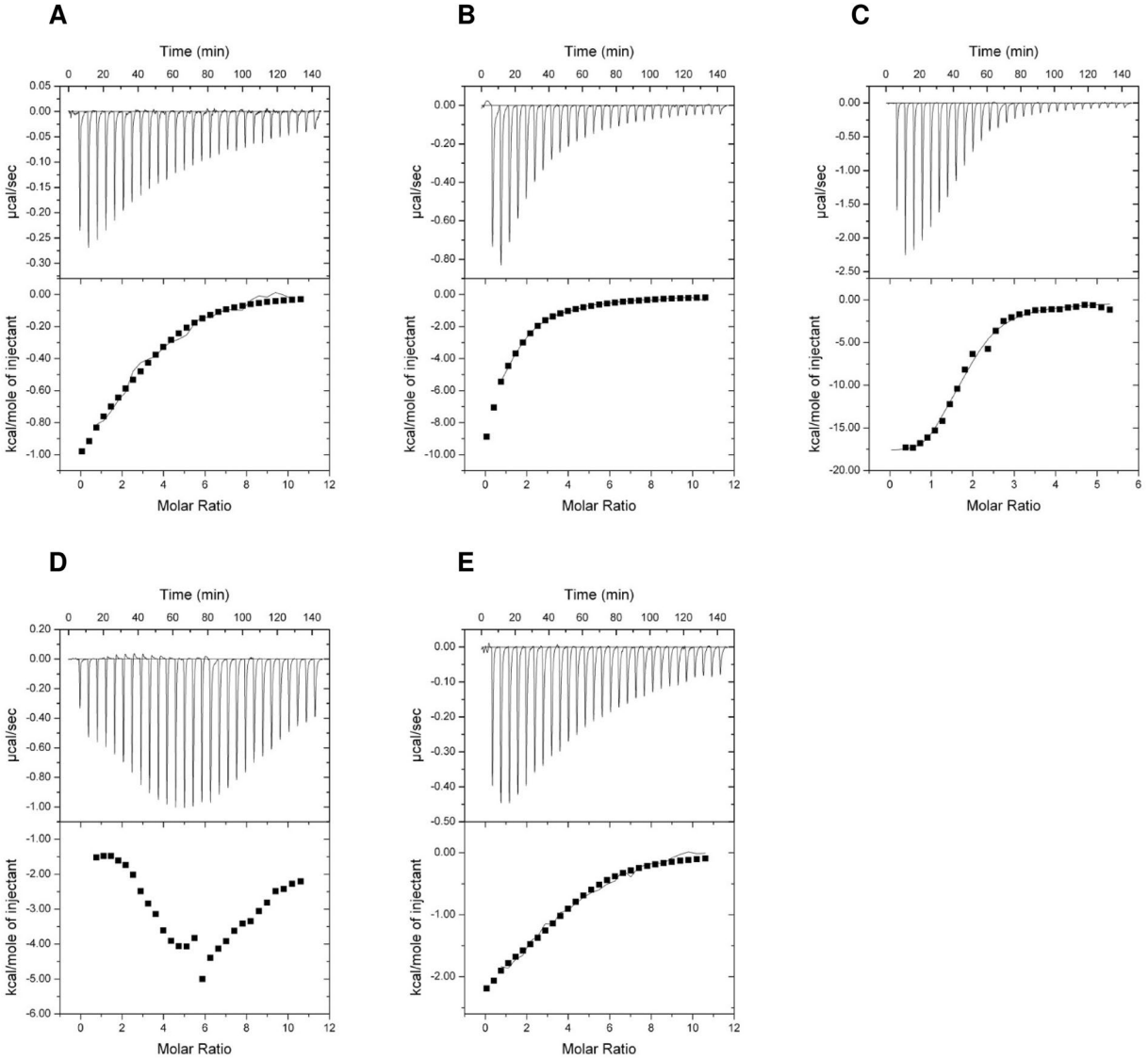


Figure 3. ITC analysis of binding small effector molecules to cyanobacterial SyPII protein. The upper panels show the raw data in the form of the heat effect during the titration of SyPII solution (trimer concentration) with ligands. The lower panels show the binding isotherm and the best-fit curve according to the three sequential binding sites model. Titration of SyPII protein (**A**) In 20 µM with 1 mM ATP. (**B**) In 20 µM with 1 mM 2-OG in the presence of 2 mM ATP. (**C**) In 20 µM with 0.5 mM ATP in presence of 1 mM 2-OG. (**D**) In 20 µM with 1 mM ADP in presence of 1 mM 2-OG. (**E**) In 20 µM with 1 mM ADP in the presence of 1 mM 2-OG and 150 µM ADP.

When PpPII in the presence of PipX was immobilized on the column, no PipX signal was detected in the elution (Supplementary Fig. S3B). This experiment showed that PpPII did not interact with PipX.

Additionally, we assessed whether PpPII could bind to PipX by surface plasmon resonance (SPR) spectroscopy (Supplementary Fig. S3C). The N-terminally His-tagged ScPipX was immobilized on a Ni-HTG sensor chip and probed with ScPII together with effector molecule ADP. Interaction between PpPII and PipX was not observed. These results supported the view that cyanobacterial PipX was not the target of PpPII.

Discussion

During the evolution of Rhodophyta, the PII proteins that are ubiquitously present in Bacteria, Archaea and in the chloroplasts of green algae and land plants[1,5,23] have been lost in most red algae. The Rhodophyta are considered to be a member of the founding lineage of photosynthetic eukaryotes (also known as Archaeplastida), whose progenitors captured the ancestral cyanobacterium-derived plastid[34,35]. Among the red algae, genes encoding putative PII proteins have only been identified until now in *P. purpurea, P. umbilicalis* and *P. yezoensis*[20,21]. In this work, PII from *P. purpurea* is characterized with respect to its functional similarities and differences compared with PII proteins from cyanobacteria and chloroplastida.

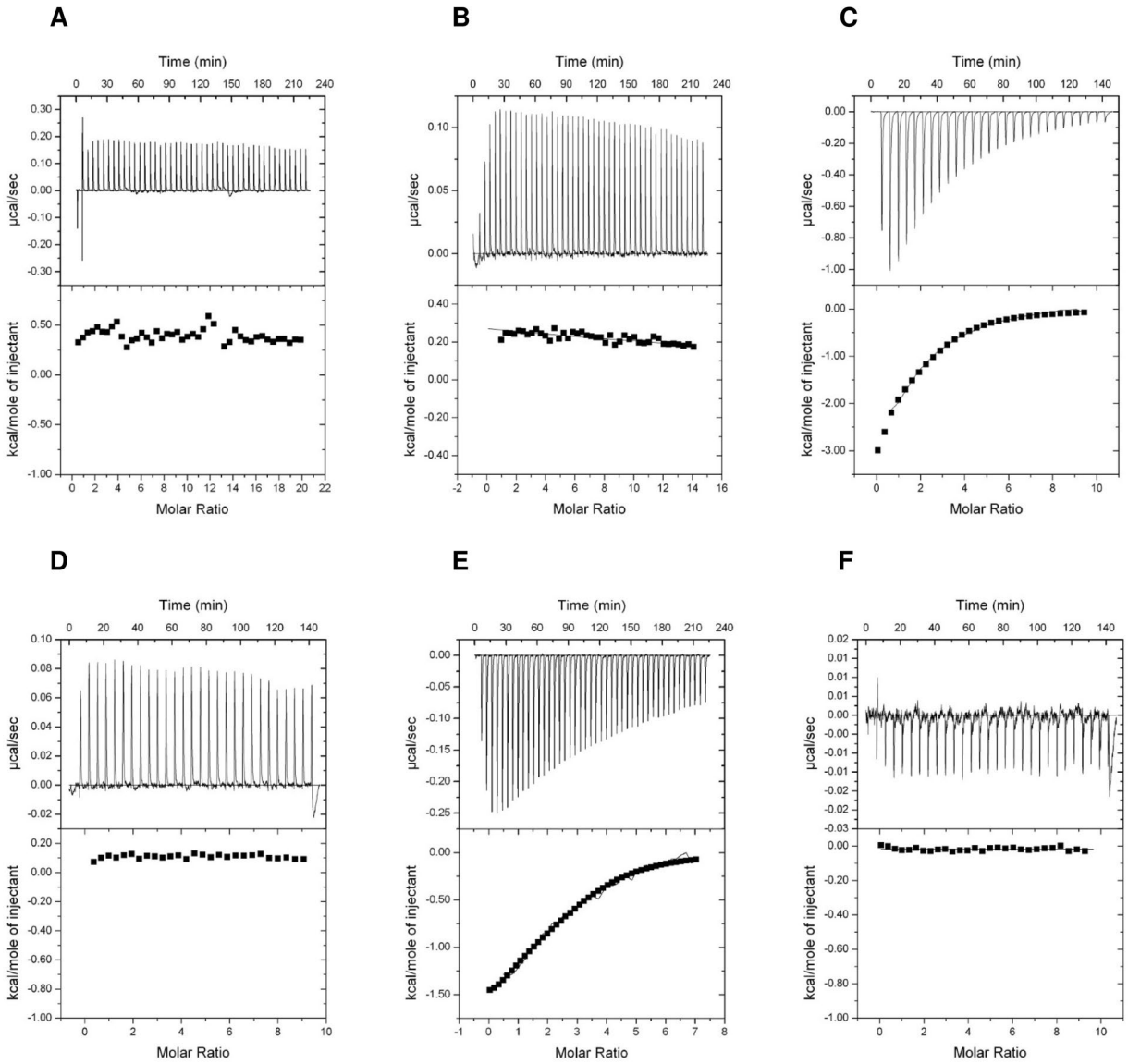


Figure 4. ITC analysis of binding small effector molecules to green algal CrPII protein. The upper panels show the raw data in the form of the heat effect during the titration of CrPII solution (trimer concentration) with ligands. The lower panels show the binding isotherm and the best-fit curve according to the three sequential binding sites model. Titration of CrPII protein (**A**) In 20 µM with 2 mM ATP. (**B**) In 28.3 µM with 2 mM ADP. (**C**) In 45 µM with 2 mM ATP in presence of 2 mM 2-OG. (**D**) In 45 µM with 2 mM ADP in presence of 2 mM 2-OG. (**E**) In 28.3 µM with 1 mM 2-OG in the presence of 2 mM ATP. (**F**) In 44 µM with 1 mM 2-OG in the presence of 2 mM ADP.

Like their bacterial PII homologs, the red algal PII proteins do not have the transit signal N- and the unique C-terminal sequences present in the green algal and land plant proteins of this family (Fig. 1). In addition, the sequences of PII from the cyanobacteria and from *P. purpurea, P. umbilicalis* and *P. yezoensis* are highly similar. All known PII proteins are able to sense and integrate signals from central metabolism: competitive binding of ADP and ATP reflects the energy state and binding of 2-OG indicates the cellular C-N balance[36–38]. The PII trimers contain three effector nucleotide-binding sites, one in each intersubunit cleft[9]. The PpPII protein of *P. purpurea* also binds three molecules of ATP and ADP, both with negative cooperativity, and with ATP having slightly higher affinity than ADP (Fig. 2A and B and Table 1), similar to the PII protein of *S. elongatus*[9]. In the presence of Mg^{2+}-ATP, the three sites can also bind 2-OG (Fig. 2E). The three 2-OG binding sites exhibit negative cooperativity to each other (Table 1) but the binding of ATP and 2-OG is synergistic to each other. This synergy of ATP and 2-OG binding is typical of bacterial PII proteins[30,36,39–41]. Conversely, 2-OG quenches binding of ADP to PpPII (Fig. 2D).

For green algal CrPII protein, we were unable to detect any binding for ATP alone, while in presence of 2-OG a strong binding isotherm for ATP was restored. This indicates that the affinity of CrPII protein towards ATP is below detection limit and 2-OG is required to constitute an efficient binding site for ATP. In the case of PhyscoPII,

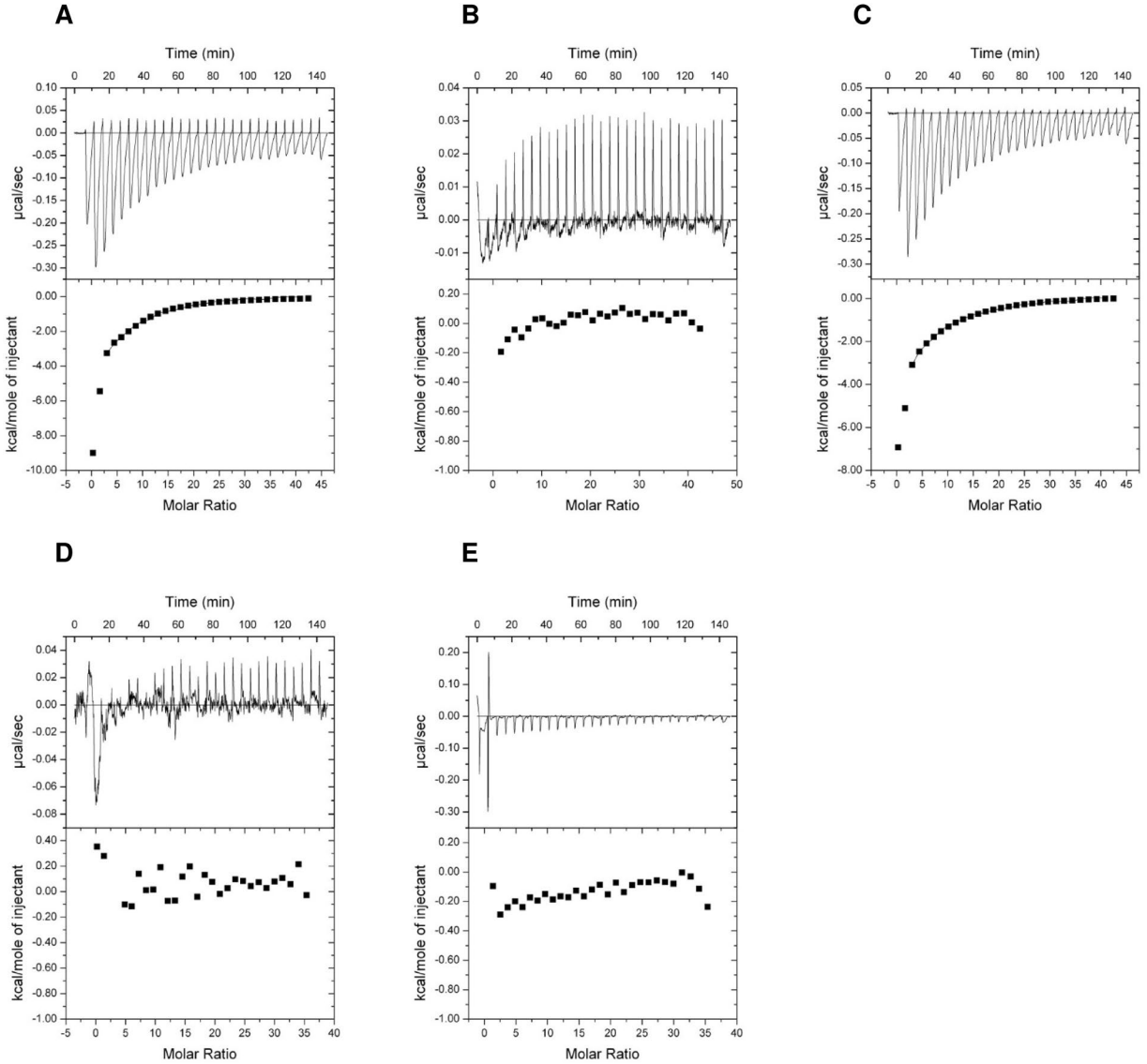


Figure 5. ITC analysis of binding small effector molecules to plant-moss PhyscoPII protein. The upper panels show the raw data in the form of the heat effect during the titration of PhyscoPII solution (trimer concentration) with ligands. The lower panels show the binding isotherm and the best-fit curve according to the three sequential binding sites model. Titration of PhyscoPII (**A**) In 5 µM with 1 mM ATP. (**B**) In 5 µM with 1 mM ADP. (**C**) In 5 µM with 1 mM ATP in presence of 2 mM 2-OG. (**D**) In 6 µM with 1 mM ADP in presence of 1 mM 2-OG. (**E**) In 6 µM with 1 mM 2-OG in the presence of 1 mM ATP.

ATP binding follows the ancestral (cyanobacterial) mode. However, we were unable to detect ADP binding for CrPII as well as for PhyscoPII protein. The loss of ADP binding suggests a differentiation in the energy sensing role of PII proteins in chloroplast evolution. Whereas in cyanobacteria and red algae, binding of ATP and ADP is competitive, the Chloroplastida seem to have specialized the PII function towards the ATP-ligated state. Furthermore, the lack of 2-OG response in PhyscoPII protein was striking. It suggests the regulatory role of 2-OG could have become obsolete in some lineages, and may be completely replaced by the later acquired glutamine sensing properties of PII proteins: green plant PII proteins are able to sense glutamine through a small C-terminal extension (termed Q-Loop) which represents a low-affinity glutamine binding site[24]. A previous publication reported that *Arabidopsis* PII protein binds 2-OG in presence on ADP[30], which contrasts the results reported here. Since *Arabidopsis* is an exception in the plant kingdom due to its secondary loss of the glutamine binding site[24], it may have compensated this loss by an enhanced affinity towards 2-OG. In some prokaryotic lineages, PII proteins were reported that lack 2-OG responses such as the protein form archaeon *Archaeoglobus fulgidus*[42].

Interestingly for ADP-binding PII proteins (red algal PpPII and cyanobacterial SyPII), the pre-equilibration of the proteins with 2-OG lowers the ADP affinities in ITC measurements, revealing strong antagonistic effect between 2-OG and ADP binding, especially in case of cyanobacterial SyPII protein, which shows complex ITC curve (Fig. 3D). The antagonistic effect of 2-OG on ADP binding affinities was mentioned previously for *A.*

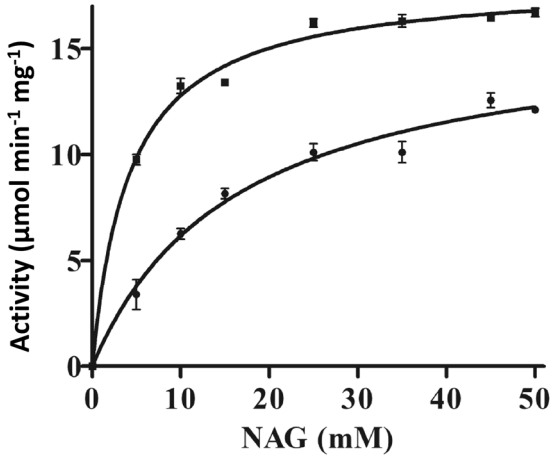


Figure 6. The catalytic activity of ScNAGK with or without PpPII protein. NAG was used as a variable substrate at fixed 10 mM ATP. The curves were fitted with GraphPad prism software. Standard deviations from triplicate experiments are indicated by error bars. (●), free ScNAGK; (■), ScNAGK with PpPII.

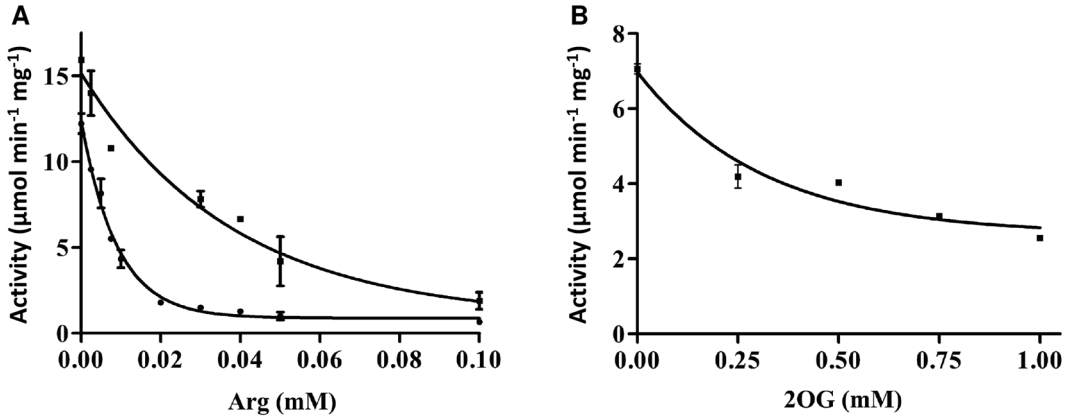


Figure 7. Arginine inhibition and antagonistic effect of 2-OG on NAGK activation by PpPII in the presence of arginine. (**A**) Inhibition of ScNAGK activity by arginine inhibition of ScNAGK in the absence (●) or presence of PpPII (■). Coupled NAGK assays were performed in the presence of 50 mM NAG and 10 mM ATP, together with increasing concentrations of arginine, as indicated. (**B**) Effect of 2-OG on PpPII activation of ScNAGK. Assays contained 50 mM NAG, 10 mM ATP, and 0.01 mM arginine, together with increasing concentrations of 2-OG, as indicated. Standard deviations from triplicate experiments are indicated by error bars.

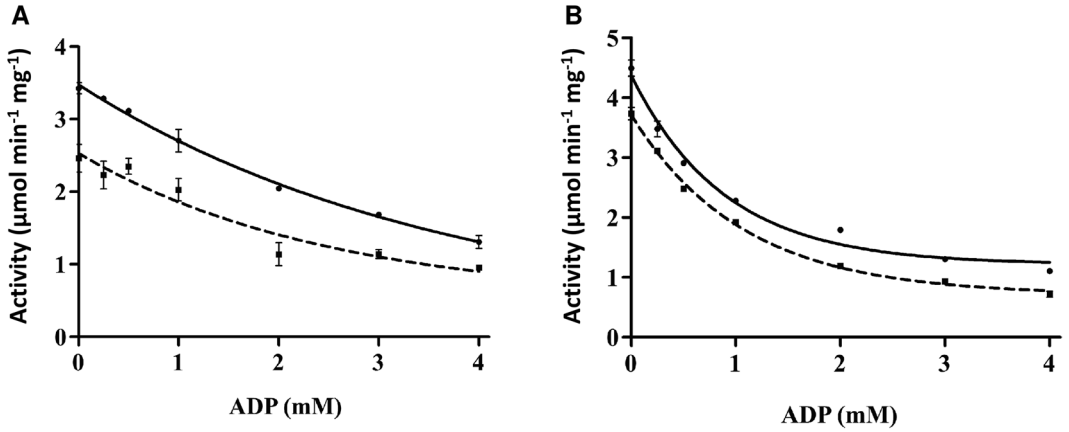


Figure 8. Effect of ADP on PpPII-mediated NAGK activity in the AGPR-coupled assay. Assays were performed in the presence of ATP at a concentration of 1 mM (dashed line) or 2 mM (continuous line) as indicated. Enzyme activity without (**A**) or with (**B**) PpPII protein. The curves were fitted with GraphPad prism software. Standard deviations from triplicate experiments are indicated by error bars.

thaliana PII[30] and explained mechanistically for *E. coli* PII protein[37]. Modeling the complex binding possibilities of PII proteins provided a mathematical interpretation of the synergetic effect between 2-OG/ATP and antagonistic effect between 2-OG/ADP. Under high physiological concentrations of 2-OG, 2-OG antagonizes ADP binding to PII by favoring the binding of the competing nucleotide ATP, since the presence of 2-OG increases the number of PII:ATP complex states[43]. Together, this demonstrates that 2-OG shifts the competition of ATP and ADP for PII binding in favor of ATP to ensure formation of the final PII:ATP:2-OG complex under high 2-OG conditions.

Altogether, the comparative analysis of PII binding properties points out that the red algal PpPII protein is quite distinct from plant PII proteins, and clearly closer to cyanobacterial PII proteins. In the blue and red lineages, 2-OG is strongly synergistic with ATP binding to PII, while antagonistic for ADP binding. In later stages of evolution of Viridiplantae, the PII proteins diverged in their properties, becoming very heterogeneous with respect to 2-OG and to ADP binding, while only the binding of ATP in presence of 2-OG has been conserved.

The interaction between PII proteins and NAGK was shown to be highly conserved from cyanobacteria to higher plants[24,25,32]. We showed previously that PII proteins across domains of life are able to interact with PII-controlled NAGK enzymes of different phyla[32]. In agreement, here we show that PpPII protein is able to interact and regulate *Synechocystis* sp. PCC 6803 NAGK (Fig. 6), which indicates the evolutionary conserved ability of the PII signaling protein to control NAGK activity across domains of life. When nitrogen is abundant, oxygenic phototrophic organisms are suggested to store nitrogen as arginine, by relieving feedback inhibition of the arginine biosynthesis controlling enzyme, NAGK[44,45]. The C-terminus of plant PII (except *Brassicaceae*) contains an extension, which binds glutamine and mediates glutamine sensing[24]. As expected from the lack of the C-terminal extension, PpPII protein relieves arginine inhibition of ScNAGK (Fig. 7A) in a glutamine-independent manner. Moreover, PpPII-mediated relief from arginine inhibition is antagonized by 2-OG (Fig. 7B).

Apparently, NAGK from cyanobacteria highly resembles NAGK from those red algae, which possess the PII protein. Intriguingly, the residues critically involved in PII interaction, are not conserved in NAGK sequences form red algae, which have lost the *glnB* gene (encoding PII) (Supplementary Fig. S2). Of particular importance is the arginine residues, which corresponds to R233 of NAGK from *S. elongatus*. This residue forms a salt bridge with B-loop residue E85 of SyPII, which represents the first step (the encounter complex) of PII-NAGK complex formation[9,46], Since the critical residues for compelx formation are conserved in PII and NAGK sequences of *P. purpurea* (Fig. 1 and Supplementary Fig. S2), we propose that complex formation between ScNAGK and red algal PpPII protein follows the same mechanism.

PII proteins from cyanobacteria directly sense the adenylate energy charge, resulting in target-dependent differential modification of the PII-signaling properties[8,31]. The kinetic activation of ScNAGK by PpPII was weakened by ADP (Fig. 8). Furthermore, in the absence of PpPII, ScNAGK responded only weakly to different ATP/ADP ratios, with about 1.8-fold reduction of activity comparing 0 and 4 mM ADP at any used ATP concentration. This indicates that the NAGK activity response towards different ATP/ADP ratios in presence of PpPII operates through competitive binding of the adenyle-nucleotides to the PII binding pocket. Together, the similarity in NAGK regulation between Rhodophyta and cyanobacteria suggests that nitrogen metabolism in red algal plastids has conserved main regulatory properties from the ancestral cyanobacterial endosymbiont. They use 2-OG and the energy state as the status reporters that control nitrogen storage metabolism.

The second well-studied interaction partner of PII from cyanobacteria is protein PipX[16,18], the transcriptional co-activator of NtcA. The inability of the PpPII protein to bind ScPipX agrees with the fact that the PipX gene was not conserved in the endosymbiontic process. Due to the loss of PipX in the endosymbiont, no selective pressure forced PpPII to maintain the PipX binding properties. Consequently, we confirmed that PipX is highly specific for cyanobacterial PII protein even with conservation of the interacting residues which required for formation of PII-PipX complex among algal PII proteins. Apparently, the transcriptional machinery of the eukaryotic host took over nitrogen-responsive gene expression and, thereby, PipX became obsolete. Identification of other PII partners in red algae as well as in cyanobacteria and plants would help to further delineate details of the endosymbotic process and to better understand the driving forces in evolution that resulted in the greening of our planet.

Materials and Methods

Cloning, expression and purification of PII proteins. A synthetic gene Block encoding *glnB* gene of PpPII with an optimized codon usage for the *E. coli* expression was synthesized by IDT, USA. The PpPII gene was amplified using the primer pair 5′-ATGAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGAAAAAGATCGAAGCTATTATTC-3′ and 5′- AAGCTTATTATTTTTCGAACTGCGGGTGGCTCCAAGCGCTAAGAGCTGCGCTGTTTAAATC-3′. The PCR products of PpPII and of *glnB* gene from *Synechocystis* sp. PCC 6803 (ScPII) were cloned into pASK-IBA3 vector (IBA, Germany) as described previously[47]. Recombinant PpPII and ScPII proteins containing a C-terminal Strep-tag II peptide were overexpressed in PII-deficient *E. coli* RB9060[48], and purified using affinity chromatography on a Strep-Tactin Superflow column (IBA) similarly as described previously for overexpression of PII protein from *S. elongatus* PCC 7942 (SyPII)[15]. For the production of recombinant PpPII with a N-terminal His_6-tag the synthetic gene Blocks fragment of PpPII was cloned directly into NdeI-digested pET15b vector (Novagen-Merck, Germany) by aqua cloning[49]. The recombinant N-terminal fused His_6-tag-PpPII protein with was overexpressed in *E. coli* Rosetta (Novagen) and purified on a Ni-NTA column as in[50]. Truncated versions of recombinant PII proteins from *P. patens* (PhyscoPII) and from *C. reinhardtii* PII (CrPII) lacking the putative transit signal peptides, starting with amino acid V60 and E63 respectively, with a C-terminally fused Strep-tag II or a N-terminally fused His_6-tag sequences were overexpressed and purified as described previously[23,24]. The quality of purified PII proteins were checked using SDS-PAGE, and the fractions containing highly purified proteins were combined (Supplementary Fig. S1) and dialyzed in the respective buffer (Supplementary Methods). The proteins concentrations were determined using Bradford reagent (Roti®-Quant, Roth). The amino acid sequences of recombinant PII proteins used in this study are described in (Supplementary Fig. S4).

Expression and purification of ScNAGK, ScPipX and AGPR proteins. The genes encoding NAGK and PipX from *Synechocystis* sp. PCC 6803 and N-acetyl-γ-glutamyl-5-phosphate reductase (AGPR) from *E. coli*, were cloned in pET15b plasmid, were induced in *E. coli* strain BL21(DE3), and the N-terminal fused His_6-tag proteins were purified using affinity chromatography on a Ni-NTA column as described[33,50].

Isothermal titration calorimetry (ITC). The ITC experiments were done using a VP-ITC microcalorimeter (MicroCal, LCC). All details for determination of binding constants for small effector molecules using ITC are fully described in the supplementary information. Heat isotherms of the dilution of the ligand in the cell buffer were collected in a blank run for each experiment in the absence of protein. The binding isotherms were calculated from received data and fitted to one-site and three-sequential binding sites models using the MicroCal ORIGIN software (Northampton) as described[9]. All titrations were performed in duplicates with different purification batches of recombinant PII proteins. The association binding constant (K_a) was generated from the software by de-convolution and curve fitting. For calculation of dissociation constant (K_d), the K_a value was inversed.

Coupled NAGK activity assay. The assess the activity of ScNAGK was measured with an enzyme assay, in which the ADP production was coupled to the NADH oxidation by the auxiliary enzymes pyruvate kinase and lactate dehydrogenase according to previously described procedure[32,50]. In a standard assay, 2.4 µg of PpPII protein was added to the reaction mix (50 mM imidazole pH 7.5, 50 mM KCl, 20 mM $MgCl_2$, 0.4 mM NADH, 1 mM phosphoenolpyruvate, 10 mM ATP, 0.5 mM DTT, 11 U lactate dehydrogenase, 15 U pyruvate kinase and 50 mM NAG) and the reaction was started by the addition of 6 µg ScNAGK. The oxidation reaction of NADH was recorded for 10 min using spectrophotometer (SPECORD 210 PLUS, Analytik Jena AG) at 340 nm. The oxidation of one molecule of NADH is proportional to the phosphorylation of one molecule of NAG. The NADH molar absorption coefficient is 6178 L mol^{-1} cm^{-1} for at 340 nm, which was used for calculation. The means of triplicates is shown with a standard deviation of less than 5%. From the velocity slopes, the enzymatic constants K_m, K_{cat} and IC_{50} were calculated using the Graph-Pad Prism software (GraphPad Software, USA).

AGPR-coupled NAGK activity assay. To assay ScNAGK activity in presence of ADP, NAGK-dependent NAG phosphorylation was coupled to the AGPR auxiliary enzyme which catalyze the reduction of NAG-phosphate using NADPH as a cofactor. The change in NADPH absorbance was recorded at 340 nm as described previously[31,33], in a reaction buffer composed of 50 mM potassium phosphate (pH 7.0), 50 mM KCl, 20 mM $MgCl_2$, 0.2 mM NADPH and 0.5 mM DTT. Each reaction contained in a volume of 1 ml 50 mM NAG, 6 μg NAGK and 10 µg of AGPR with/without 2.4 µg of PpPII. The reaction was started by the addition of ScNAGK and was recorded for 10 min with a spectrophotometer (SPECORD 210 PLUS, Analytik Jena). Kinetic constants were calculated as descried above.

Pull-down analysis. For complex formation, the binding reactions were done in 300 µl of buffer (100 mM Tris HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA) by adding 10 µM of purified proteins: His_6-ScPipX and Strep-tag ScPII or Strep-tag PpPII. The proteins were mixed in the presence of 1 mM ADP, kept at room temperature for 30 min, and then loaded on StrepTactin spin columns. Unbound proteins were removed by washing ten times per 50 µl by centrifugation with the binding buffer for 30 s, 800 rpm. (MiniSpin, Eppendorf) and bound proteins were eluted in binding buffer contains 10 mM d-desthiobiotin (Sigma-Aldrich). The eluted fractions were analyzed by Tricine-SDS PAGE[51] and stained with SimplyBlue™ SafeStain (Invitrogen).

Surface plasmon resonance (SRP) analysis. SPR experiments were conducted using a ProteOn XPR36 system (Bio-Rad) in HBS buffer (10 mM HEPES, 150 mM NaCl, 1 mM $MgCl_2$, and 0.005% Tween 20, pH 7.5) at 25 °C. Immobilization of ligand, 1 µM His_6-PipX (200 µl), was performed on a Ni^+-loaded HTG sensor chip (Bio-Rad Laboratories) in the vertical orientation into the channel L2, and the continuous running buffer was used at a flow rate of 30 µl/min. The immobilization level was approximately 3000 resonance units (RU). Next, the chip was rotated at 90°, the channels were washed for 3 min with buffer HBS, and 200 nM (200 µl) of strep-tagged PII proteins (ScPII or PpPII) was simultaneously injected in the vertical orientation into different channels at a flow rate of 30 µl/min for an association phase of 400 s, which was followed by a 600 s dissociation phase. Channel L1, which was treated with the buffer without protein, served as a reference. The binding of PII proteins to PipX was recorded as a response signal difference of L2 and L1. All binding sensorgrams were collected, processed and analyzed using the integrated ProteOn Manager software (Bio-Rad Laboratories).

References

1. Sant'Anna, F. H. *et al.* The PII superfamily revised: a novel group and evolutionary insights. *J. Mol. Evol.* **68**, 322–336 (2009).
2. Huergo, L. F., Chandra, G. & Merrick, M. PII signal transduction proteins: nitrogen regulation and beyond. *FEMS Microbiol. Rev.* **37**, 251–283 (2013).
3. Chellamuthu, V. R., Alva, V. & Forchhammer, K. From cyanobacteria to plants: conservation of PII functions during plastid evolution. *Planta.* **237**, 451–462 (2013).
4. Arcondéguy, T., Jack, R., Merrick & M. P. (II) signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol. Mol. Biol. Rev.* **65**, 80–105 (2001).
5. Forchhammer, K. PII signal transducers: novel functional and structural insights. *Trends Microbiol.* **16**, 65–72 (2008).
6. Uhrig, R. G., Ng, K. K. & Moorhead, G. B. PII in higher plants: a modern role for an ancient protein. *Trends Plant Sci.* **14**, 505–511 (2009).
7. Forchhammer, K. & Lüddecke, J. Sensory properties of the PII signalling protein family. *FEBS J.* **283**, 425–437 (2016).
8. Lüddecke, J. & Forchhammer, K. Energy sensing versus 2-oxoglutarate dependent ATPase switch in the control of *Synechococcus* PII interaction with its targets NAGK and PipX. *PLoS One.* **10**(8), e0137114 (2015).
9. Fokina, O., Chellamuthu, V. R., Zeth, K. & Forchhammer, K. A novel signal transduction protein P(II) variant from *Synechococcus elongatus* PCC 7942 indicates a two-step process for NAGK-P(II) complex formation. *J. Mol. Biol.* **399**, 410–421 (2010).

10. Truan, D. *et al*. A new PII protein structure identifies the 2-oxoglutarate binding site. *J. Mol. Biol.* **400**, 531–539 (2010).
11. Radchenko, M. & Merrick, M. The role of effector molecules in signal transduction by PII proteins. *Biochem. Soc. Trans.* **39**, 189–194 (2011).
12. Zeth, K., Fokina, O. & Forchhammer, K. An engineered PII protein variant sensing a novel ligand: atomic resolution structure of the complex with citrate. *Acta Crystallogr. Sect. D.* **68**, 901–908 (2012).
13. Ninfa, A. J. & Atkinson, M. R. PII signal transduction proteins. *Trends Microbiol.* **8**, 172–179 (2000).
14. Leigh, J. A. & Dodsworth, J. A. Nitrogen regulation in bacteria and archaea. *Ann. Rev. Microbiol.* **61**, 349–377 (2007).
15. Heinrich, A., Maheswaran, M., Ruppert, U. & Forchhammer, K. The *Synechococcus elongatus* P_{II} signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. *Mol. Microbiol.* **52**, 1303–1314 (2004).
16. Espinosa, J., Forchhammer, K., Burillo, S. & Contreras, A. Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. *Mol. Microbiol.* **61**, 457–469 (2006).
17. Espinosa, J., Forchhammer, K. & Contreras, A. Role of the *Synechococcus* PCC 7942 nitrogen regulator protein PipX in NtcA-controlled processes. *Microbiology.* **153**, 711–718 (2007).
18. Llácer, J. L. *et al*. Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. *Proc. Natl. Acad. Sci. USA* **107**, 15397–15402 (2010).
19. Hauf, W., Schmid, K., Gerhardt, E. C., Huergo, L. F. & Forchhammer, K. Interaction of the nitrogen regulatory protein GlnB (PII) with biotin carboxyl carrier protein (BCCP) controls acetyl-CoA levels in the cyanobacterium *Synechocystis* sp. PCC 6803. *Front Microbiol.* **7**, 1700 (2016).
20. Chan, C. X. *et al*. *Porphyra* (Bangiophyceae) transcriptomes provide insights into red algal development and metabolism. *J. Phycol.* **48**, 1328–1342 (2012).
21. Nakamura, Y. *et al*. The first symbiont-free genome sequence of marine red alga, Susabi-nori (*Pyropia yezoensis*). *PLoS One.* **8**(3), e57122 (2013).
22. Hsieh, M. H., Lam, H. M., van de Loo, F. J. & Coruzzi, G. A PII-like protein in Arabidopsis: putative role in nitrogen sensing. *Proc. Natl. Acad. Sci. USA* **95**, 13965–13970 (1998).
23. Ermilova, E. *et al*. PII signal transduction protein in *Chlamydomonas reinhardtii*: localization and expression pattern. *Protist.* **164**, 49–59 (2013).
24. Chellamuthu, V. R. *et al*. A widespread glutamine sensing mechanism in the plant kingdom. *Cell.* **159**, 1188–1199 (2014).
25. Minaeva, E., Forchhammer, K. & Ermilova, E. Glutamine assimilation and feedback regulation of l-acetyl-n-glutamate kinase activity in *Chlorella variabilis* NC64A results in changes in arginine pools. *Protist.* **166**, 493–505 (2015).
26. Chan, C. X. *et al*. Analysis of *Porphyra* membrane transporters demonstrates gene transfer among photosynthetic eukaryotes and numerous sodium-coupled transport systems. *Plant Physiol.* **158**, 2001–2012 (2012).
27. Blouin, N. A., Brodie, J. A., Grossman, A. C., Xu, P. & Brawley, S. H. *Porphyra*: a marine crop shaped by stress. *Trends Plant Sci.* **16**, 29–37 (2011).
28. Stiller, J. W. *et al*. Major developmental regulators and their expression in two closely related species of *Porphyra* (Rhodophyta). *J. Phycol.* **48**, 883–896 (2012).
29. Oliveira, M. A. *et al*. 2-Oxoglutarate levels control adenosine nucleotide binding by *Herbaspirillum seropedicae* PII proteins. *FEBS J.* **282**, 4797–809 (2015).
30. Smith, C. S., Weljie, A. M. & Moorhead, G. B. Molecular properties of the putative nitrogen sensor PII from *Arabidopsis thaliana*. *Plant J.* **33**, 353–360 (2003).
31. Fokina, O., Herrmann, C. & Forchhammer, K. Signal-transduction protein P(II) from *Synechococcus elongatus* PCC 7942 senses low adenylate energy charge *in vitro*. *Biochem. J.* **440**, 147–156 (2011).
32. Beez, S., Fokina, O., Herrmann, C. & Forchhammer, K. N-acetyl-Lglutamate kinase (NAGK) from oxygenic phototrophs: PII signal transduction across domains of life reveals novel insights in NAGK control. *J. Mol. Biol.* **389**, 748–758 (2009).
33. Takahara, K., Akashi, K. & Yokota, A. Continuous spectrophotometric assays for three regulatory enzymes of the arginine biosynthetic pathway. *Anal. Biochem.* **368**, 138–147 (2007).
34. Chan, C. X. *et al*. Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. *Current Biol.* **21**, 328–333 (2011).
35. Price, D. C. *et al*. *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science.* **335**, 843–847 (2012).
36. Fokina, O., Chellamuthu, V. R., Forchhammer, K. & Zeth, K. Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* PII signal transduction protein. *Proc. Natl. Acad. Sci. USA* **107**, 19760–19765 (2010).
37. Jiang, P. & Ninfa, A. J. Sensation and signaling of a-ketoglutarate and adenylylate energy charge by the *Escherichia coli* PII signal transduction protein require cooperation of the three ligand-binding sites within the PII trimer. *Biochemistry.* **48**, 11522–11531 (2009).
38. Zeth, K., Fokina, O. & Forchhammer, K. Structural basis and target specific modulation of ADP sensing by the *Synechococcus elongatus* PII signaling protein. *J. Biological Chem.* **289**, 8960–8972 (2014).
39. Jiang, P. & Ninfa, A. J. *Escherichia coli* PII signal transduction protein controlling nitrogen assimilation acts as a sensor of adenylate energy charge *in vitro*. *Biochemistry.* **46**, 12979–12996 (2007).
40. Radchenko, M. V., Thornton, J. & Merrick, M. Control of AmtB-GlnK complex formation by intracellular levels of ATP, ADP, and 2-oxoglutarate. *J. Biol. Chem.* **285**, 31037–31045 (2010).
41. Truan, D., Bjelic, S., Li, X. D. & Winkler, F. K. Structure and thermodynamics of effector molecule binding to the nitrogen signal transduction PII protein GlnZ from *Azospirillum brasilense*. *J. Mol. Biol.* **426**, 2783–2799 (2014).
42. Helfmann, S., Lü, W., Litz, C. & Andrade, S. L. Cooperative binding of MgATP and MgADP in the trimeric P(II) protein GlnK2 from *Archaeoglobus fulgidus*. *J. Mol. Biol.* **402**, 165–177 (2010).
43. da Rocha, R. A. *et al*. Mathematical model of the binding of allosteric effectors to the *Escherichia coli* PII signal transduction protein GlnB. *Biochemistry.* **52**, 2683–2693 (2013).
44. Caldovic, L. & Tuchman, M. N-acetylglutamate and its changing role through evolution. *Biochem. J.* **372**, 279–290 (2003).
45. Llácer, J. L., Fita, I. & Rubio, V. Arginine and nitrogen storage. *Curr. Opin. Struc. Biol.* **18**, 673–681 (2008).
46. Llácer, J. L. *et al*. The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. *Proc. Natl. Acad. Sci. USA* **104**, 17644–17649 (2007).
47. Gibson, D. G. *et al*. Enzymatic assembly of DNA moleculesup to several hundred kilobases. *Nat. Methods.* **6**, 343–345 (2009).
48. Bueno, R., Pahel, G. & Magasanik, B. Role of glnB and glnD gene products in regulation of the glnALG operon of *Escherichia coli*. *J. Bacteriol.* **164**, 816–822 (1985).
49. Beyer, H. M. *et al*. Aqua cloning: a versatile and simple enzyme-free cloning approach. *PLoS One.* **10**(9), e0137652 (2015).
50. Maheswaran, M., Urbanke, C. & Forchhammer, K. Complex formation and catalytic activation by the PII signaling protein of N-acetyl-L-glutamate kinase from *Synechococcus elongatus* strain PCC 7942. *J. Biol. Chem.* **279**, 55202–55210 (2004).
51. Schägger, H., Tricine-SDS-PAGE. *Nat. Protoc.* **1**, 16–22 (2006).

Acknowledgements

This work was supported by Saint-Petersburg State University (research Grant No. 1.65.38.2017) and RFBR (N15-54-12370) to E.E., and DFG to K.F. (Fo195/9-2, Fo195/13-1). K.A.S. received a scholarship from the

Deutsche Akademische Austauschdienst (DAAD). SPR analysis was performed at the research resource center for molecular and cell technologies of Saint-Petersburg State University. All authors would like to thank Jörg Scholl (Organismic Interactions Department, Tübingen University) for providing the cyanobacterial plasmids used in this study.

Author Contributions

E.E., K.A.S. and K.F. designed the study and wrote the manuscript. T.L. purified algal and plant proteins, characterized NAGK enzyme activity and performed pull-down and SRP analyses. K.A.S. purified cyanobacterial proteins and performed all ITC experiments. All authors analyzed the results and approved the final version of the manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-19046-7.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

P_{II}-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response

Khaled A. Selim[a,b,1], Florian Haase[c], Marcus D. Hartmann[b], Martin Hagemann[c], and Karl Forchhammer[a,1]

[a]Organismic Interactions Department, Interfaculty Institute for Microbiology and Infection Medicine, Tübingen University, 72076 Tübingen, Germany; [b]Department of Protein Evolution, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany; and [c]Plant Physiology Department, Institute of Biological Sciences, Rostock University, 18059 Rostock, Germany

Edited by Susan S. Golden, University of California, San Diego, La Jolla, CA, and approved April 9, 2018 (received for review March 7, 2018)

Cyanobacteria are phototrophic prokaryotes that evolved oxygenic photosynthesis ∼2.7 billion y ago and are presently responsible for ∼10% of total global photosynthetic production. To cope with the evolutionary pressure of dropping ambient CO_2 concentrations, they evolved a CO_2-concentrating mechanism (CCM) to augment intracellular inorganic carbon (C_i) levels for efficient CO_2 fixation. However, how cyanobacteria sense the fluctuation in C_i is poorly understood. Here we present biochemical, structural, and physiological insights into SbtB, a unique P_{II}-like signaling protein, which provides new insights into C_i sensing. SbtB is highly conserved in cyanobacteria and is coexpressed with CCM genes. The SbtB protein from the cyanobacterium *Synechocystis* sp. PCC 6803 bound a variety of adenosine nucleotides, including the second messenger cAMP. Cocrystal structures unraveled the individual binding modes of trimeric SbtB with AMP and cAMP. The nucleotide-binding pocket is located between the subunit clefts of SbtB, perfectly matching the structure of canonical P_{II} proteins. This clearly indicates that proteins of the P_{II} superfamily arose from a common ancestor, whose structurally conserved nucleotide-binding pocket has evolved to sense different adenyl nucleotides for various signaling functions. Moreover, we provide physiological and biochemical evidence for the involvement of SbtB in C_i acclimation. Collectively, our results suggest that SbtB acts as a C_i sensor protein via cAMP binding, highlighting an evolutionarily conserved role for cAMP in signaling the cellular carbon status.

P_{II}-like protein SbtB | signal transduction | inorganic carbon signaling | cAMP | cyanobacteria

Cyanobacteria evolved oxygenic photosynthesis about 2.7 billion y ago. This ability was later transmitted to eukaryotes via endosymbiosis, giving rise to plastids in algae and plants (1). The global proliferation of oxygenic photosynthetic conversion of C_i (CO_2 and HCO_3^-) into organic matter using light energy and electrons from water splitting led over time to a strong decline in available C_i on Earth. As a consequence, cyanobacteria evolved a CO_2-concentrating mechanism (CCM) that allows them to grow in the present-day atmosphere of the Earth, which contains only 0.04% CO_2 (2, 3).

The cyanobacterial CCM consists of several systems for C_i uptake as well as of the carboxysome, a proteinaceous bacterial cellular compartment in which the CO_2-fixing enzyme RubisCO and carbonic anhydrase (CA) are localized. The C_i-uptake systems concentrate massive amounts of bicarbonate inside the cyanobacterial cells, which then diffuses into the carboxysome, where CA converts it to CO_2. This strategy increases the CO_2 concentration in the vicinity of RubisCO, thereby saturating the carboxylation activity and minimizing the competing oxygenation reaction of this enzyme (2, 3). Therefore, the ability to respond to fluctuating C_i supply is of key importance, but the underlying mechanisms are still poorly understood.

The cyanobacterial bicarbonate transporter SbtA (4), as part of the CCM, is encoded in a bicistronic operon, together with the downstream gene *sbtB*; this operon is highly expressed with other CCM genes under C_i-limiting conditions (5, 6). In *Escherichia coli* cells expressing cyanobacterial *sbtA* and *sbtB*, SbtB inhibits SbtA-mediated HCO_3^- uptake (7). SbtB has been identified as a noncanonical member of the P_{II} signal transduction superfamily (8). P_{II} proteins are among the most widespread signaling proteins in nature. Canonical P_{II} proteins regulate nitrogen assimilation reactions by sensing the metabolic state of the cells through binding of small effector molecules [2-oxoglutarate (2-OG) and ATP/ADP]. This induces conformational changes in the large T loop that changes P_{II} interactions with target proteins, thereby modulating their activities. However, although members of the P_{II} superfamily are highly conserved in structure, they are very divergent in sequence, such that P_{II}-like proteins usually have low sequence identity and lack multiple consensus sequence motifs of canonical P_{II} proteins (8). The cognate effector molecules and biological functions of most P_{II}-like proteins, such as SbtB, are unknown. Here we analyzed the sensing properties and regulatory features of SbtB from *Synechocystis* sp. PCC 6803 (*Sc*SbtB).

Results

Biochemical Properties of *Sc*SbtB. Analysis of the *Sc*SbtB amino acid sequence (*SI Appendix*, *Bioinformatics Analysis* and Fig. S1 *A* and *B*) revealed the absence of virtually all residues known to

Significance

Life on Earth depends on photosynthetic CO_2 fixation to form organic carbon. This process evolved in cyanobacteria and was later conveyed to eukaryotes, giving rise to plastids in algae and plants. To cope with low atmospheric CO_2 concentrations that developed over the course of evolution, cyanobacteria evolved a CO_2-concentrating mechanism (CCM), which elevates CO_2 levels in the vicinity of RubisCO, the key enzyme of CO_2 fixation. Here we describe a conserved cyclic AMP receptor protein, SbtB, which participates in the sensing of fluctuating C_i levels to regulate the cyanobacterial CCM system. SbtB represents a new principle of C_i sensing, which is important for acclimation to varying C_i regimes in the ecological niches of cyanobacteria.

Author contributions: K.A.S., M.D.H., M.H., and K.F. designed research; K.A.S. performed biochemical, structural, and biophysical experiments; K.A.S. and F.H. performed physiological experiments; K.A.S., M.D.H., M.H., and K.F. analyzed data; and K.A.S., M.D.H., M.H., and K.F. wrote the manuscript.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Crystallography, atomic coordinates, and structure factors have been deposited in the Protein Data Bank, www.wwpdb.org (PDB ID codes 5O3P, 5O3Q, 5O3R, and 5O3S).

[1]To whom correspondence may be addressed. Email: khaled.selim@uni-tuebingen.de or karl.forchhammer@uni-tuebingen.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1803790115/-/DCSupplemental.

Published online May 7, 2018.

MICROBIOLOGY

interact with 2-OG or the β- and γ-phosphates of the effector nucleotides (ATP/ADP) of canonical P_{II} proteins, which suggests that SbtB evolved sensory properties distinct from those of canonical P_{II} proteins. To identify potential sensing properties and regulatory features of SbtB, we prepared a recombinant *Sc*SbtB protein in its native trimeric state (*SI Appendix*, Fig. S1 *C–E*) and set out to address the question of potential metabolite binding using isothermal titration calorimetry (ITC) and microscale thermophoresis (MST).

We first tested the binding of the typical P_{II} effectors ATP, ADP, and 2-OG. While 2-OG binding was not detected (*SI Appendix*, Fig. S2), both ATP and ADP bound with comparably high affinities in the micromolar range (Table 1), even though SbtB lacks the canonical core residues for β- and γ-phosphate coordination. This led us to test for the binding of adenosine and AMP, which represent the core moiety for which large parts of the binding site appeared to be conserved. Binding of adenosine was weak (*SI Appendix*, Fig. S2), and AMP bound with an affinity only slightly lower than that of ATP and ADP, which suggested that AMP is the minimal binding moiety that is specifically recognized (*SI Appendix*, Fig. S2 *A–C*). We then checked for the binding of cAMP, which surprisingly bound with the highest affinity of all compounds tested (*SI Appendix*, Fig. S2*D*). Moreover, the binding enthalpy for cAMP was higher than that of the other nucleotides at the same concentration (compare *SI Appendix*, Fig. S2*E* with *SI Appendix*, Fig. S2 *A–C*), which indicated that binding cAMP is preferred over the other nucleotides. The guanosine nucleotides GTP and cGMP did not bind to *Sc*SbtB (*SI Appendix*, Fig. S2). For all nucleotides, a comparison of the calculated affinities using a model of three binding sites showed that the first two sites have similar affinities, whereas the third site is occupied with an approximately 10-fold lower affinity, which is indicative of anticooperativity (Table 1), like the canonical P_{II} protein (9, 10). The K_d values obtained by the MST measurements were in good qualitative agreement with the ITC data (Fig. 1*A* and Table 1).

The recent characterization of the P_{II}-like protein associated with carboxysomes from *Thiomonas intermedia* (*Ti*CP_{II}), whose gene is located near the *sbtAB* operon in many chemoautotrophic bacteria, provided clues that bicarbonate, AMP, and ADP are ligands (11). In addition, the binding of the adenosine nucleotides to *Ti*CP_{II} is stimulated 15-fold by 50 mM bicarbonate (11). Therefore, we also tested the effect of bicarbonate on *Sc*SbtB. Bicarbonate (50 mM) did not bind to *Sc*SbtB (*SI Appendix*, Fig. S2) or significantly influence the binding properties of the verified ligands (Table 1). To ensure that binding of cAMP is unique to *Sc*SbtB and not an overlooked property of canonical P_{II} proteins, we titrated cAMP and AMP against the *Synechococcus elongatus* P_{II} protein (*Se*P_{II}) without detecting any binding (*SI Appendix*, Fig. S2 *L* and *M*).

Structural Characterization of *Sc*SbtB. To determine the structural basis for ligand binding, *Sc*SbtB in apo form and in complex with AMP, ADP, or cAMP was crystallized, all rewarding well-diffracting crystals. The apo protein yielded a trigonal crystal form of space group $P3_2$. We solved its structure by molecular replacement using the *Anabaena variabilis SbtB* structure [*Av*SbtB; Protein Data Bank (PDB) ID code 3DFE] as a search model (*SI Appendix*, *Structural Characterization of* Sc*SbtB*). As expected, the refined apo-*Sc*Sbtb structure shows a canonical P_{II} ferredoxin-like fold (Fig. 1 *B* and *C*). All three subunits of the trimer are essentially in the same conformation, and the entire *Sc*SbtB trimer superimposes with Cα rmsd values of 0.53 Å on the *Av*SbtB trimer and 0.93 Å on the canonical *Se*P_{II} trimer (PDB ID code 2XUL) (*SI Appendix*, Figs. S3 and S4*A*). As in essentially all other crystal structures of isolated trimers of the P_{II} superfamily, the flexible T loop was mostly disordered (Fig. 1*C*).

Notable overall differences from the structures of canonical P_{II} proteins are found at the N and C termini. At the N terminus, the first β-strand runs antiparallel along the full length of the β4-strand to face the C terminus (Fig. 1*C*). The C terminus features a unique extension, which was unresolved in the *Av*SbtB structure but is fully ordered here. It forms a single α-helical turn flanking the canonical nucleotide-binding cleft, followed by a short hairpin loop formed by a CGPxGC motif, which is conserved in SbtB proteins (*SI Appendix*, Fig. S1 *A* and *B*), in which the two cysteines form a disulfide bridge (Fig. 1*C*). The carboxyl terminus of the last cysteine forms a hydrogen bond with the amino terminus. This mutually stabilizing interaction between the *Sc*SbtB N and C termini is shown in detail in Fig. 1*C*. The whole structural motif is invariant in all three chains of the trimer, which underpins its relevance.

Structural Characterization of Ligand Binding. Cocrystals of *Sc*SbtB with AMP, ADP, or cAMP yielded the same trigonal crystal form as for apo-*Sc*SbtB, and the same crystals were obtained in the presence and absence of bicarbonate. We solved the structures

Table 1. Dissociation constants calculated for binding of the adenyl nucleotides ATP, ADP, AMP, cAMP, and adenosine to the SbtB protein measured by ITC and MST

	ITC			MST
Titrant/Protein	K_{d1}, μM	K_{d2}, μM	K_{d3}, μM	K_d, μM
ATP				
SbtB	22.0 ± 1.0	6.9 ± 1.5	158.6 ± 23.8	46.0 ± 7.0
SbtB + 50 mM HCO_3^-	11.7 ± 0.2	5.6 ± 0.8	103.9 ± 1.9	NA
ADP				
SbtB	7.6 ± 1.0	9.3 ± 3.5	150.6 ± 18.8	18.9 ± 3.7
SbtB + 50 mM HCO_3^-	9.1 ± 2.3	7.2 ± 0.04	127.1 ± 13.4	NA
AMP				
SbtB	10.1 ± 1.1	7.4 ± 0.3	244.9 ± 9.8	70.4 ± 7.9
SbtB + 50 mM HCO_3^-	29.9 ± 33.8	16.1 ± 13.5	173.6 ± 20.3	NA
cAMP				
SbtB	2.1 ± 0.6	12.6 ± 2.5	76.2 ± 11.9	11.0 ± 1.9
SbtB + 50 mM HCO_3^-	3.9 ± 0.4	6.2 ± 1.2	86.8 ± 28.4	NA
Adenosine/SbtB		Weak binding		1,800.0 ± 1,530.0

The ITC K_d values were assayed in the presence or absence of 50 mM HCO_3^-. The raw ITC data were fitted to a model of three sequential binding sites for trimeric *Sc*SbtB. The K_d values correspond to the mean of the independent experiments ± SD. NA, not tested.

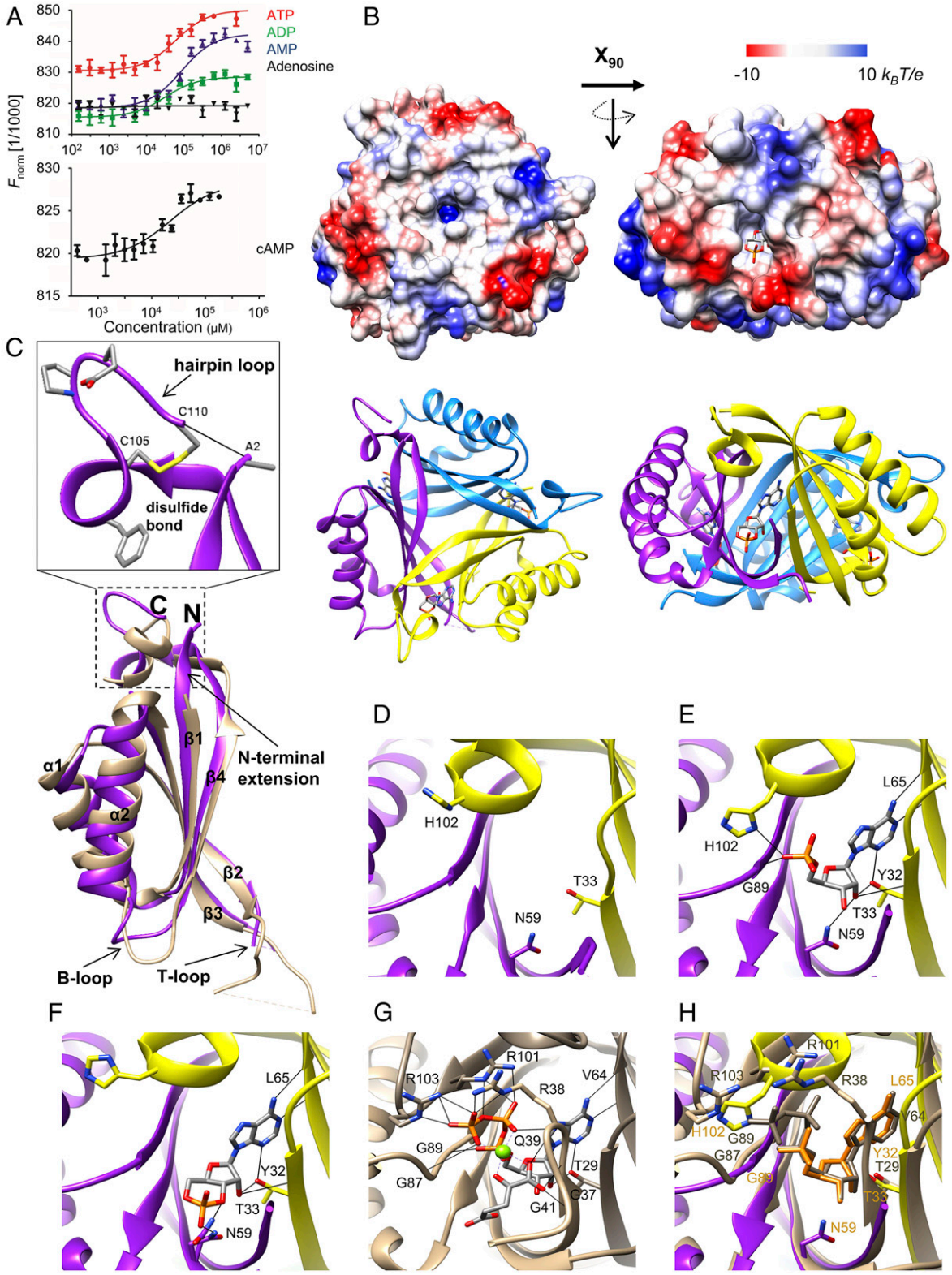


Fig. 1. Structural and binding properties of the *Sc*SbtB protein. (*A*) Binding of nucleotides monitored by MST ($n = 3$). (*A*, *Upper*) Titration series of ATP, ADP, AMP, and adenosine. (*A*, *Lower*) Titration series of cAMP. (*B*) Overall architecture of the SbtB protein. The structure of the trimeric SbtB–cAMP complex is shown as an electrostatic potential surface (*Upper*) and as a cartoon (*Lower*). The red–white–blue color gradient corresponds to surface potential values ranging from −10 to +10 k_BT/e. The structure of SbtB reveals the typical ferredoxin-like fold of the P_{II} superfamily, with nucleotide-binding pockets situated between the clefts of the subunits. (*C*) Superposition of monomeric subunits of SbtB (violet) with *Se*P_{II} (brown; PDB ID code 2XUL), yielding a 0.93-Å rmsd, with secondary structure elements and characteristic structural motifs indicated. (*C*, *Inset*) Highlighting of the hairpin-loop (CGPEGC) motif of SbtB with the disulfide bond between C105 and C110 and the hydrogen bond to the N terminus. (*D*–*G*) Close-up of the nucleotide binding site of the apo-SbtB, SbtB–AMP, SbtB–cAMP, and *Se*P_{II}–ATP structures, respectively. Relevant residues for nucleotide binding are shown as sticks, and H bonds are indicated by thin lines. (*H*) Superposition of the SbtB–AMP (yellow–violet) and *Se*P_{II}–ATP (brown) nucleotide binding sites, with relevant residues labeled in orange and dark brown, respectively.

MICROBIOLOGY

on the basis of the apo-*Sc*SbtB coordinates; in all cocrystal structures, the electron density of the respective nucleotide was clear (*SI Appendix*, Fig. S4*B*). Structures of crystals grown in the presence or absence of bicarbonate did not differ, in agreement with the negligible influence of bicarbonate on the ITC measurements, which further confirms that the sensory properties of *Sc*SbtB differ from those of *Ti*CP$_{II}$ (11).

To our surprise, however, not only in the *Sc*SbtB–AMP structure but also in the putative *Sc*SbtB–ADP structure, only AMP was complexed, with clearly no electron density for a β-phosphate. Consequently, these two structures are virtually identical. The AMP moiety is bound as in most other known nucleotide-bound structures of the P$_{II}$ superfamily (Fig. 1 *D–H*). Specifically, the adenosine moiety forms hydrogen bonds to the backbone of L65 and the T33 side chain, and the phosphate forms a hydrogen bond to G89. Additional *Sc*SbtB-specific hydrogen bonds are formed between the ribose hydroxyl groups and the side chain of N59, which is only conserved in *Sc*SbtB, in the β-strand (β3) following the T loop. In two chains of the trimer, another bond is observed between the phosphate and a histidine (H102) in the C extension (Fig. 1*E*).

Due to the cyclic nature of cAMP, its phosphate binding mode differs from that of AMP. While the interactions of the adenine and ribose moieties of cAMP are the same as in the *Sc*SbtB–AMP structure, no bonds are formed between the phosphate and G89 or H102. Instead, at least in one binding site, the phosphate forms an additional hydrogen bond with N59 (Fig. 1*F*). All residues involved in nucleotide binding are shown in detail in *SI Appendix*, Fig. S4*B*.

ScSbtB Reversibly Associates with Membranes in an SbtA-Dependent Manner. *Sc*SbtB-specific antibodies (*SI Appendix*, Fig. S5*A*) were raised to reveal the subcellular localization and expression pattern of *Sc*SbtB in response to different treatments, showing that *Sc*SbtB accumulates in stationary-phase and carbon-limited conditions (*SI Appendix*, *Conditions Triggering SbtB Accumulation* and Fig. S5). It has been reported that when *sbtB* and *sbtA* are coexpressed in *E. coli*, SbtB tunes down SbtA-dependent HCO_3^- uptake (7). If the two proteins interact in wild-type *Synechocystis* sp. PCC 6803 cells (hereinafter WT), then SbtB should colocalize with the integral membrane protein SbtA. In WT cultures grown with low carbon (ambient air; hereinafter LC), SbtB was equally distributed between the membrane and soluble (cytosolic) fractions (Fig. 2*A*, *Left*). When cells were exposed to harsh C$_i$ starvation by omitting any gas exchange for 24 h (stationary culture flasks without aeration), SbtB was mainly in the membrane fraction (Fig. 2*A*, *Right*). By contrast, in an SbtA-deficient mutant, SbtB was exclusively soluble and never membrane-associated (Fig. 2*B*), which indicates that SbtA is required for SbtB membrane localization. To determine the dynamics of SbtB association/dissociation to/from the membrane, a WT culture grown with high carbon (5% CO_2; hereinafter HC) was shifted to LC conditions. At the start of the experiment (T_{zero}), SbtB was detected exclusively in the soluble fraction. One hour after the shift, the first traces of membrane-localized SbtB were observed, and the membrane-bound SbtB fraction gradually increased over 6 h (Fig. 2*C*).

To determine whether the ligands cAMP, AMP, and ADP affect the membrane localization of SbtB, each ligand (2 mM) was separately added to extracts of WT cells grown with LC and the extracts were then fractionated into soluble and membrane fractions. In WT extracts without added ligand, SbtB was partially localized in the membrane. Addition of ADP or AMP clearly and markedly triggered membrane localization of SbtB, whereas cAMP did not cause membrane localization of SbtB (Fig. 2*D*). Hence, binding of ADP or AMP, in contrast to binding of cAMP, would be advantageous for SbtA–SbtB complex formation in the membrane.

We then created a cAMP-free mutant by knocking out *slr1991* (designated *cya1*), which encodes the soluble adenylyl cyclase (12) (*SI Appendix*, Fig. S6 *A* and *B*), and tested the effect on SbtB localization upon shifting from LC to HC conditions. Under LC conditions, the amount of SbtB associated with the membrane was similar in the Δ*cya1* mutant and the WT. However, after shifting to HC conditions, SbtB dissociated from the membrane in the WT but not in the Δ*cya1* mutant (Fig. 2*E*). This result indicates that SbtB in the Δ*cya1* mutant does not sense the HC signal, which suggests that cAMP represents the HC signal that leads to membrane dissociation of SbtB.

ScSbtB Is Required for Efficient C$_i$ Acclimation. To characterize the role of SbtB in the CCM, we generated a completely segregated *sbtB* knockout mutant (Δ*sbtB*) (*SI Appendix*, Figs. S5*A* and S6 *A* and *C*). On solid BG$_{11}$ agar medium at pH 8 and 9, the Δ*sbtB* mutant grew almost like the WT under HC/LC conditions, but growth was negatively affected only at pH 7 under LC conditions, where more C$_i$ is present in the form of CO_2 (Fig. 3*A*, *Left*). Similarly, the Δ*sbtB* mutant grew significantly slower than the WT in liquid cultures at pH 8 under LC conditions, but such a

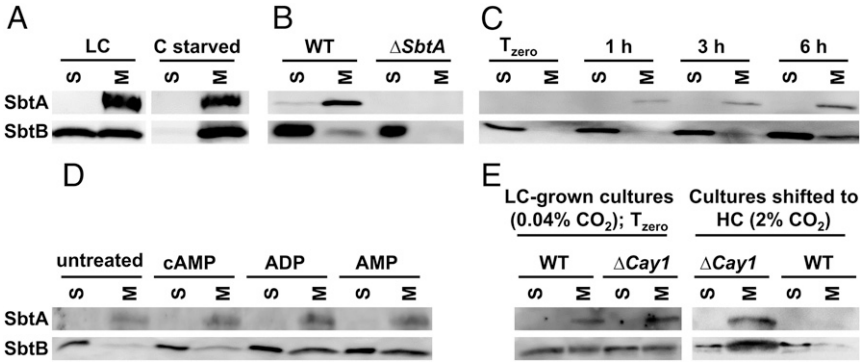


Fig. 2. Immunoblot analysis of SbtA-dependent *Sc*SbtB membrane localization. (*A*) Localization of SbtB in soluble (S) and membrane (M) fractions of cells grown with ambient air, that is, LC (0.04% CO_2; *Left*), or in carbon-starved cells (stationary culture flasks without aeration for 24 h; *Right*). (*B*) Distribution of SbtB between soluble and membrane fractions of WT and Δ*sbtA* mutant cells grown with LC (0.04% CO_2). SbtA antibodies were used to determine the quality of membrane isolation and the presence of SbtA. (*C*) Dynamics of SbtB membrane localization in WT cultures shifted from HC (5% CO_2) to LC (0.04% CO_2) for 1 to 6 h; T_{zero} represents HC conditions before the shift. (*D*) Influence of added nucleotides on SbtB localization. Crude cell extracts of cells grown under LC were prepared, the indicated nucleotide (2 mM) was then added separately to the extracts, and membrane and soluble fractions were separated. (*E*) SbtB localization in WT and Δ*cya1* mutant cells grown under LC (air-grown) and HC (2% CO_2) conditions as indicated. See *SI Appendix*, *Materials and Methods* for a detailed description of membrane fractionation.

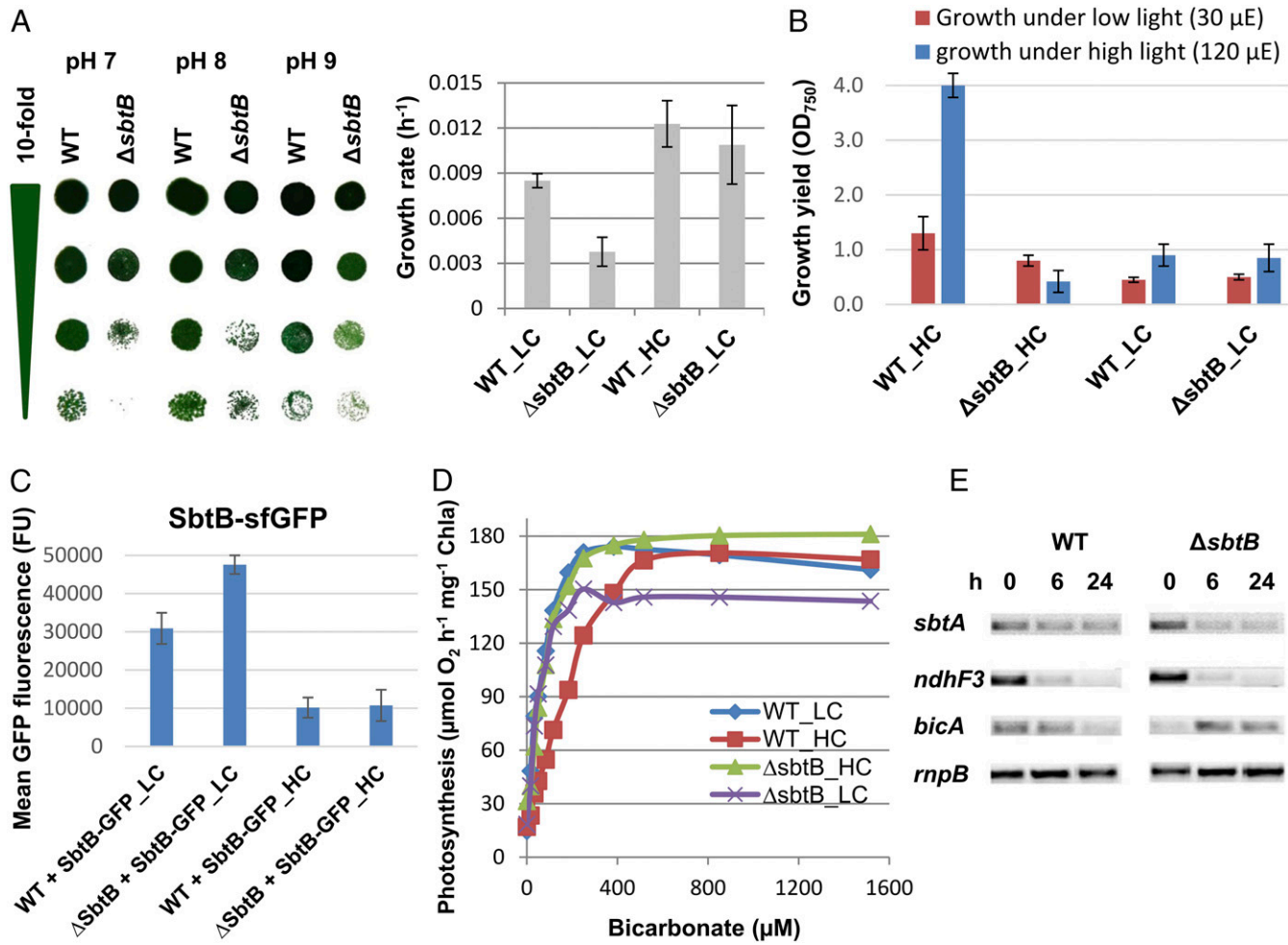


Fig. 3. Phenotypic characterization of the Δ*sbtB* knockout mutant. (*A*) Growth of WT or Δ*sbtB* mutant cells on solid BG_{11} medium at different pH values under LC conditions (*Left*) or in liquid BG_{11} medium at pH 8.0 under LC or HC conditions (*Right*, error bars represent standard deviations from three independent cultivation experiments). (*B*) Effect of sudden change in carbon supply in the form of growth yield. WT and the Δ*sbtB* mutant were shifted from slow-growing cultures (LC/low light, 30 μE) either to HC (2% CO_2) or LC (air) under the initial light intensity (30 μE), or to HC (5% CO_2) and LC (air) with increased light intensity (120 μE). (*C*) Levels of fluorescent SbtB-sfGFP protein in WT and the Δ*sbtB* complemented mutant in response to HC and LC growth (standard deviation; $n = 3$). (*D*) Bicarbonate-dependent photosynthetic rates per chlorophyll a (Chla) of WT and the Δ*sbtB* mutant as a function of increasing HCO_3^- concentrations. Cells were acclimated to either HC or LC conditions ($n = 3$). (*E*) Gene expression of selected CCM genes in WT and the Δ*sbtB* mutant analyzed using semiquantitative RT-PCR. Cells were cultivated under LC conditions (0 h) and then shifted to HC conditions (5% CO_2) for 24 h; the constitutively expressed *rnpB* gene served as loading control.

difference was not observed under HC conditions (Fig. 3*A*, *Right*). Importantly, Δ*sbtB* cells were highly sensitive to sudden changes in carbon supply; when slowly growing Δ*sbtB* mutant cells (LC and low light of 30 μE) were suddenly exposed to 2% CO_2 at the same light intensity, the culture temporarily ceased growth and only slowly recovered. This effect became even more pronounced at higher light intensity and with further CO_2 supplementation (5% CO_2, 120 μE). By contrast, WT cultures immediately increased growth at HC levels under low and high light intensity (Fig. 3*B* and *SI Appendix*, *Physiological Function of* Sc*SbtB* and Fig. S7*A*).

To complement the growth defect of Δ*sbtB*, a plasmid expressing the gene encoding an *Sc*SbtB-sfGFP (superfolded GFP) fusion protein under the control of the native promoter of the *sbt* operon was transformed into WT and Δ*sbtB* cells to generate the strains WT+SbtB-sfGFP and Δ*sbtB*+SbtB-sfGFP (*SI Appendix*, Fig. S6 *A* and *D*). To check the expression of the SbtB-sfGFP construct, we quantified SbtB-sfGFP fluorescence under HC and LC conditions (Fig. 3*C*). The GFP signal was much lower under HC conditions than under LC conditions, in agreement with the expression pattern of native SbtB in response to C_i (5, 6). Under LC, the SbtB-sfGFP signal was lower in the WT+SbtB-sfGFP strain than in the Δ*sbtB*+SbtB-sfGFP strain, which indicated that the genomic copies of *sbtB* (*slr1513*) in WT cells quench the expression of SbtB-sfGFP. Next, we tested whether the *Sc*SbtB-sfGFP construct complements the phenotype of the Δ*sbtB* mutant by analyzing acclimation to a shift in CO_2 levels as described above. Indeed, the SbtB-sfGFP construct complemented the growth defect of the Δ*sbtB* mutant (*SI Appendix*, Fig. S7 *A–D*), which demonstrated that SbtB-sfGFP was functionally active and that the growth defect of Δ*sbtB* was not caused by polar effects. Last, we determined the cellular localization of SbtB under HC and LC conditions by monitoring the sfGFP signals by fluorescence microscopy. Under HC conditions, the SbtB-sfGFP signal was mainly cytoplasmic. However, 1 h after the shift to LC conditions, SbtB-sfGFP partially relocalized to the cytoplasmic membrane (*SI Appendix*, Fig. S8). Together, these data suggest that SbtB is required to acclimate to sudden changes in CO_2 levels, which highlights its sensory/regulatory role in C_i acclimation.

To obtain further evidence for the proposed functional link between SbtB and control of C_i metabolism, we determined C_i-uptake

MICROBIOLOGY

affinity from HC- and LC-acclimated cells by measuring the amount of oxygen evolved from HCO_3^--dependent photosynthesis. Intriguingly, the maximal photosynthetic rates of LC-acclimated Δ*sbtB* mutant cells were lower than those of the WT (Fig. 3*D*). The lower V_{max} indicates an impaired function of CO_2-fixing Calvin–Benson cycle activity, which limits the photosynthetic activity at saturating C_i amounts; this result is consistent with the slower growth of Δ*sbtB* under LC conditions. In contrast, the affinity toward bicarbonate was almost similar in WT and mutant Δ*sbtB* cells under LC conditions, since the initial rise in the photosynthetic activity was indistinguishable between these two strains. Strikingly, only WT cells showed the expected acclimation to HC conditions by lowering bicarbonate affinity (HCO_3^- K_m: LC, 48 μM; HC, 95 μM), whereas Δ*sbtB* cells retained the high C_i affinity under HC conditions (HCO_3^- K_m: LC, 52 μM; HC, 68 μM) (Fig. 3*D*). This behavior implies that Δ*sbtB* resides in the LC-adapted state even under HC conditions, and provides a rationale for the inability of the Δ*sbtB* mutant to rapidly acclimate to a CO_2 upshift (see above).

The inability of the Δ*sbtB* mutant to adjust the C_i affinity in response to the ambient CO_2 supply suggests that the SbtB protein plays a role in the regulation of the entire CCM, similar to the role of the canonical P_{II} protein in N assimilation (8, 13). To test whether the Δ*sbtB* mutation affects the expression of other C_i-controlled genes, we analyzed the transcript levels of CCM-related genes upon shifting cells from LC to HC. The genes coding for various C_i-uptake systems were highly expressed in WT and Δ*sbtB* mutant cells grown under LC conditions (Fig. 3*E*, time 0), and expression of *sbtA* and *ndhF3* markedly decreased upon shifting to HC conditions; *ndhF3* encodes a subunit of the inducible CO_2-uptake system NDH-I_3 (3, 14) (Fig. 3*E*, time points 6 and 24 h after shift). Since the abundance of *sbtA* mRNA was similar in the WT and Δ*sbtB* mutant, we can safely assume that the downstream *sbtB* gene mutation does not affect the expression of the upstream *sbtA* gene; this result agrees with a previous report that the inactivation of *sbtB* in *Synechococcus* sp. PCC 7942 does not affect the bicarbonate-uptake activity of SbtA (15). Surprisingly, the expression of another HCO_3^- transporter, *bicA* (16), was deregulated completely in Δ*sbtB*, whereas in the WT the expression was like that of *sbtA* (Fig. 3*E*), which indicates that the *sbtB* mutation might have a broader impact on C_i regulation than only the direct regulation of SbtA. The expression of the LC-inducible gene *bicA* under HC conditions agrees with the preacclimatization of the Δ*sbtB* mutant to LC conditions, and might be the cause of the inability of the Δ*sbtB* mutant to adapt C_i affinity to HC conditions.

cAMP/AMP as a C_i Signal Compete for *Sc*SbtB. To address whether the stably bound *Sc*SbtB effector molecules cAMP and AMP are related to fluctuations in C_i supply, we measured the cellular levels of cAMP and AMP in WT cells as a function of C_i. An HC-adapted culture was harvested, resuspended in fresh medium, and divided into two subcultures. One subculture was reincubated again under HC conditions, and the other subculture was incubated under LC conditions. After 30 min and 4 h, samples were taken, and the levels of cAMP and AMP were quantified. The AMP levels were 18- to 20-fold higher under LC conditions than under HC conditions. By contrast, cAMP showed the opposite trend, namely cAMP levels were 3- to 4-fold higher under HC conditions than under LC conditions (Fig. 4*A*).

To reveal whether these physiological changes in cAMP relative to AMP are functionally relevant for *Sc*SbtB signaling, we determined the competitive binding of cAMP and AMP to *Sc*SbtB using ITC. Under saturating concentrations of cAMP (300 μM), binding of AMP (2 mM) to *Sc*SbtB was completely inhibited (Fig. 4*B*). In the presence of 100 μM cAMP, AMP was able to bind only to the available unoccupied sites of *Sc*SbtB with low affinity (*SI Appendix*, Fig. S2*N*). The determined K_d value of 240.9 μM is similar to the third (low-affinity) binding site of AMP in trimeric *Sc*SbtB (compare with Table 1). However, when *Sc*SbtB was

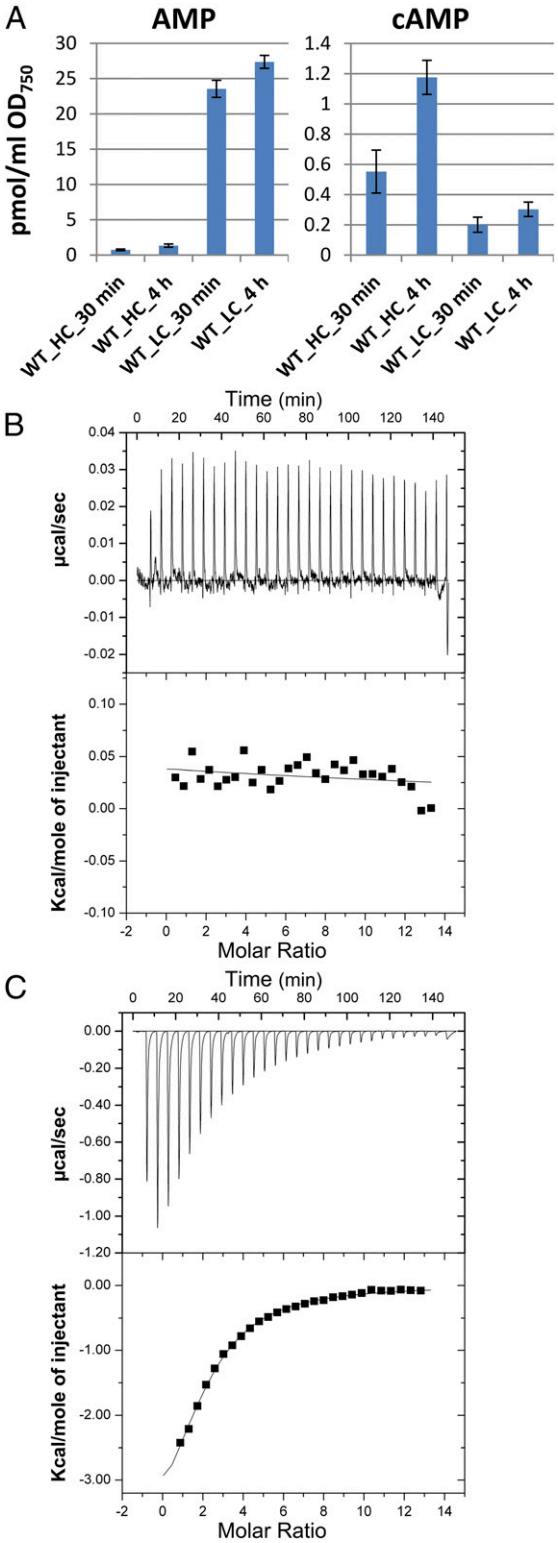


Fig. 4. Response of cAMP and AMP to cellular C_i levels, and competitive binding assay of cAMP and AMP to *Sc*SbtB. (*A*) Mean cellular levels and standard deviations ($n = 3$) of cAMP and AMP were determined under LC or HC conditions after 30 min and 4 h. (*B* and *C*) ITC titration of 33.3 μM SbtB (trimeric concentration) in the presence of (*B*) 300 μM cAMP against 2 mM AMP and (*C*) 300 μM AMP against 2 mM cAMP. (*B* and *C*, *Upper*) Raw ITC data in the form of the heat produced during the titration. (*B* and *C*, *Lower*) Binding isotherms and the best-fit curves according to the one–binding-site model for trimeric SbtB.

preincubated with 300 μM AMP, cAMP was still able to bind to *Sc*SbtB, with each injection (corresponding to ∼14 μM cAMP) producing a strong exothermic signal. The resulting isotherm could be fitted to an apparent K_d of 42.2 μM (Fig. 4*C*). Therefore, under HC conditions, with elevated levels of cAMP and highly diminished levels of AMP, we expect that *Sc*SbtB will accommodate cAMP. By contrast, under LC conditions, the increasing levels of AMP and concomitantly decreasing levels of cAMP will give rise to the AMP-bound state of *Sc*SbtB.

Discussion

Common properties of canonical P_{II} proteins are the competitive binding of the adenyl nucleotides ATP or ADP to sense the energy status of cells (8, 17) and the ATP-dependent binding of 2-OG as a status indicator of the cellular carbon–nitrogen balance. By contrast, the larger family of P_{II}-like proteins has different adenyl-nucleotide binding modes and does not generally bind 2-OG (8, 18). SbtB extends the range of sensing in the P_{II} superfamily. *Sc*SbtB is a unique member of the P_{II} superfamily for which specific binding of ATP, ADP, AMP, and cAMP has been described. In contrast to canonical PII proteins, the β- and γ-phosphates of the adenyl nucleotides seem of minor importance for SbtB binding and, accordingly, the interacting residues, such as the RxR motif in the C loop, are not conserved (*SI Appendix*, Fig. S1). However, the location of the nucleotide-binding pocket in the intersubunit clefts and the recognition mode of the nucleoside moiety are perfectly conserved with other members of the P_{II} superfamily (8) (Fig. 1*H*). This agrees with the assumption that proteins of the P_{II} superfamily arose from a common nucleotide-binding ancestor, whose binding pocket has been structurally conserved but whose binding mode has evolved to sense different adenyl nucleotides to fulfill various signaling functions (8).

Interestingly, when we solved the structure of the putative *Sc*SbtB–ADP complex, AMP was bound and there was clearly no electron density for a β-phosphate. A similar phenomenon has been reported for other members of the P_{II} superfamily (11, 19), but whether ADP is enzymatically hydrolyzed by *Sc*SbtB to AMP remains unclear.

An intriguing property of *Sc*SbtB is the presence of a conserved CGPxGC motif at the C terminus. This extension forms a small helix–hairpin loop structure, which forms a disulfide bond between cysteines C105 and C110. An identical C-terminal segment, CGPxGC, is also present in a subgroup of thiol disulfide oxidoreductases (DsbA). The members of this subgroup form dimers in which the CGPxGC terminus of one monomer is in close proximity to the catalytic CPxC motif. The structure of such a C terminus of a DsbA family protein (PDB ID code 3GL5) is almost identical to the hairpin-loop extension of *Sc*SbtB. Furthermore, many of the plant Calvin–Benson–cycle enzymes have evolved similar redox-regulated C-terminal extensions, C(V/I)VxVC, to modulate their activity through oxidation/reduction of intramolecular disulfide bridges under the different light conditions of day/night cycles (20, 21). Therefore, the hairpin-loop extension of SbtB might be involved in a redox-sensory function.

One of the most striking features of *Sc*SbtB is the binding of cAMP. Increasing evidence suggests that cAMP signaling plays an important role in C_i perception. The increased levels of cAMP under HC conditions agree with the finding that cyanobacterial soluble adenylyl cyclase activity is stimulated by elevated levels of C_i, which causes increased cAMP production (22–25). In diatoms, which possess CCMs, increasing CO_2 concentrations cause elevated cAMP levels, which induce down-regulation of the CCMs (26). Therefore, cAMP can be considered an evolutionarily conserved second messenger molecule that responds to C_i, thereby regulating CCMs in deeply branched phyla of cyanobacteria and algae. Another well-known regulatory nucleotide, the stringent control nucleotide ppGpp, was recently identified in cyanobacteria to control global gene expression during day/night switches (27, 28), emphasizing a fundamental role of secondary messenger nucleotides in the cyanobacteria lifestyle.

Although it has been recognized that cAMP is involved in carbon sensing, our results imply that the level of AMP might also play a role in C_i sensing. The levels of cAMP and AMP were similar in cells grown under HC conditions, whereas the cAMP-to-AMP ratio decreased ∼60- to 80-fold when cultures were shifted to LC conditions (Fig. 4). From this, we expect that under HC conditions, SbtB will preferentially bind cAMP over AMP, whereas under LC conditions, AMP will outcompete cAMP. In vitro, AMP favored the membrane-associated state of SbtB, whereas cAMP favored the soluble state. This perfectly agrees with the prevalence of cytoplasmic SbtB under HC conditions (high cAMP-to-AMP ratio) as opposed to membrane-associated SbtB in the LC state (low cAMP-to-AMP ratio). Structurally, there is no considerable difference between the AMP- and cAMP-bound SbtB conformations. However, it is conceivable that the bound nucleotides directly contribute to the interface between SbtA and SbtB, such that the open phosphate group of AMP stabilizes the complex, while cAMP disrupts the interface with its differently oriented phosphate.

*Sc*SbtB and the recently described *Ti*CP_{II} protein (PDB ID code 5DS7) both sense ADP and AMP (11), and superimpose closely with an rmsd of 0.9 Å (*SI Appendix*, Figs. S3*C* and S4*A*). However, *Ti*CP_{II} directly responds to HCO_3^- and thereby increases the affinity to the adenyl nucleotides (11), whereas *Sc*SbtB does not respond to HCO_3^-, its affinity to adenyl nucleotides is not influenced by high amounts of external bicarbonate levels (Table 1), and it senses cAMP as a second messenger of the cellular carbon supply. Since both P_{II}-like proteins respond to ADP/AMP, it appears that a central, yet not understood, task of these proteins is to link C_i assimilation to the energy and carbon states of the cells. However, the sensing of the C_i status occurs via distinct mechanisms. In the case of *Ti*CP_{II}, HCO_3^- is directly bound in the effector molecule binding site, thereby increasing the affinity by two orders of magnitude. By contrast, *Sc*SbtB perceives the carbon status indirectly, via the binding of the second messenger cAMP.

The phenotypic changes of the SbtB-deficient mutant suggest that SbtB is involved in the regulation of the cyanobacterial CCM, similar to the role of the canonical P_{II} protein in N assimilation (8, 13). Cells of the SbtB-deficient mutant seem to be locked in the LC-acclimated state even under HC conditions. Gene expression analysis revealed that the *sbtB* deletion had no impact on HC-induced repression of the upstream *sbtA* gene, whereas expression of the *bicA* gene encoding another bicarbonate transporter differed in WT and mutant cells. Currently, we have no mechanistic explanation for this finding, but it indicates that different factors are involved in the regulation of *bicA* and the *sbt* operon, with *bicA* expression being dependent on SbtB. Therefore, in addition to a direct regulation of SbtA transport activity (7), SbtB might be involved in a more general regulation of the CCM, in analogy to the multiple functions of canonical cyanobacterial P_{II} proteins in metabolic regulation (29).

Under conditions of C_i starvation, SbtA and SbtB were produced at the highest levels, and SbtB was bound to the membrane via SbtA. Under these conditions, maximal activity of SbtA is required to keep the intracellular bicarbonate level high. Consequently, this SbtA–SbtB complex must be active in bicarbonate transport and, therefore, binding of SbtB does not necessarily inhibit SbtA activity. This clearly distinguishes SbtB from the P_{II} protein GlnK, which associates with the ammonium channel protein Amt to close the pores and shut down the uptake of ammonia (30, 31). It is, however, possible that SbtB switches off SbtA activity during the night, possibly involving the C terminus of *Sc*SbtB, which is likely redox-regulated. Another purpose of SbtB binding to SbtA could be to stabilize SbtA or to sequester SbtB to the membrane and thereby exert a regulatory function, similar to that shown for the regulation of nitrogenase-associated DraG/DraT by

MICROBIOLOGY

P_{II} proteins in *Azospirillum brasilense* (32). If soluble SbtB interacts with a regulator that mediates acclimation to high CO_2 levels, then deletion of SbtB could lock the cells in the LC-adapted state.

Collectively, our findings extend the knowledge on cAMP sensing. While the sensing of organic carbon (glucose) by cAMP is well-understood, a mechanistic link for the thus far unclear relationship between cAMP and C_i sensing has been enigmatic. The P_{II}-like signaling protein SbtB represents a novel type of cAMP receptor protein and perceives C_i status via binding of the second messenger cAMP. In eukaryotes, cAMP transmits signals that originate at G protein-coupled receptors and modulate cellular activities mainly through cAMP-dependent protein kinases (33), thereby affecting gene expression or central metabolism, in particular glycogen, sugar, and lipid metabolism. In bacteria, cAMP mediates manifold cellular responses, mainly via cAMP-responsive transcription factors, such as CRP (34, 35), but also via cAMP-controlled protein modification enzymes, such as protein kinases and GNAT-related protein lysine acetylases (36). With SbtB, a novel mechanism of cAMP-mediated signaling in cyanobacterial carbon regulation has been unraveled, and a functional link between CO_2 acquisition and cAMP sensing has been revealed. However, further studies are required to mechanistically understand the interaction with the target SbtA as well as the interplay between the various adenyl-nucleotide effector molecules for energy and carbon sensing.

Materials and Methods

Full protocols are available in *SI Appendix, Materials and Methods*.

Purification of Recombinant Proteins. Recombinant C-terminal StrepII-tagged proteins (*Sc*SbtB and canonical *Se*P_{II}) were purified as previously described (9, 10).

Determination of *Sc*SbtB Binding Constants. *Sc*SbtB binding constants were determined using ITC (VP-ITC microcalorimeter; MicroCal) and MST (Monolith NT.LabelFree; NanoTemper Technologies). The ITC binding isotherms were calculated from the recorded data and fitted to one-site and three-site binding models using MicroCal Origin software (10). MST single-site fitting was done using NanoTemper data analysis software.

Crystallization, Data Collection, and Structure Elucidation. Proteins were crystallized via vapor diffusion at 20 °C in 96-well sitting-drop plates. The structure of apo-*Sc*SbtB was solved by molecular replacement using MOLREP (37) and the *Av*SbtB structure (PDB ID code 3DFE) as a search model. Other structures were solved subsequently on the basis of the apo-*Sc*SbtB coordinates (PDB ID code 5O3P). Structural representations were prepared using UCSF Chimera (https://www.cgl.ucsf.edu/chimera/).

Generation of Mutants. Knockout deletion mutants were constructed by cloning the upstream and downstream regions of the ORF *slr1513* (designated *sbtB*) (4, 5) and the ORF *slr1991* (designated *cya1*) (12) with the erythromycin resistance cassette (primers are listed in *SI Appendix*, Table S1). For complementation, SbtB-sfGFP strains were generated by introducing the *sbtB* gene (*slr1513*) fused to the gene encoding sfGFP into Δ*sbtB* and WT backgrounds using the self-replicating plasmid pVZ322.

CO_2-Shift Experiments. CO_2-shift experiments were performed according to ref. 38. Briefly, cells were precultured in standard BG_{11} medium (pH 8) with bubbling of air-enriched CO_2 [2 to 5% (vol/vol); defined as high-carbon conditions] at a light intensity of about 100 μE. Cells were shifted to low CO_2 by bubbling with ambient air [0.04% (vol/vol) CO_2; defined as low-carbon conditions]. For a harsh carbon shift, cells were transferred to BG_{11} medium free of HCO_3^- and CO_3^{-2} and kept stationary. For a shift to HC conditions, the gas stream of LC-grown cells was switched from ambient air to CO_2-supplemented air.

Phenotypic Characterization of the Δ*sbtB* Mutant. Wild-type *Synechocystis* sp. PCC 6803 and Δ*sbtB* cells were grown under various CO_2 conditions as mentioned above. The survival of cells was checked in standard BG_{11} medium in a drop-dilution assay on agar plates of various pH and by measuring growth rates in liquid medium. The rate of C_i-dependent oxygen evolution as a function of increasing HCO_3^- supply was determined using a Clark-type oxygen electrode (Hansatech).

Transcriptional Analysis. Total RNA was isolated using the PGTX method (39). Reverse transcription was done using the RevertAid RT Kit (Thermo Fisher). Aliquots of the generated cDNA were used for the RT-PCR assays of selected C_i-regulated genes using gene-specific primers (*SI Appendix*, Table S1).

Quantification of cAMP and AMP. cAMP and AMP were extracted according to ref. 40. The intracellular levels of cAMP and AMP were analyzed and determined using the LCMS-8050 System and LC/MS/MS Method Package for Primary Metabolites version 2 (Shimadzu). For quantification of cAMP and AMP, authentic substances in appropriate concentrations were used for calibration.

Immunoblot Analysis of the Cellular Localization of *Sc*SbtB. For immunoblot analysis, the crude cell extract from WT cells grown in or shifted to different CO_2 conditions as indicated was prepared using a ribolyser for four cycles at 4 °C. For localization experiments, total cell extracts were fractionated into membranes and soluble fractions according to ref. 30.

ACKNOWLEDGMENTS. The authors are grateful to Reinhard Albrecht for crystallographic sample preparation and assistance with diffraction data collection, the staff of beamline X10SA/SLS, and Karen Brune for critical linguistic editing of the manuscript. M.D.H. gratefully acknowledges Andrei Lupas for continued support. K.A.S. and K.F. thank all members of the Organismic Interactions Department, especially Jan Lüddecke and Carolina Nishi, for technical assistance. This work was supported by Deutsche Forschungsgemeinschaft Grant Fo195/9-2, Graduiertenkolleg Grant 1708-2, and Deutsch Akademischer Austauschdienst (DAAD); the Hochschulbauförderungsgesetz (HBFG) program financed the LC/MS equipment (GZ: INST 264/125-1 FUGG).

1. Hohmann-Marriott MF, Blankenship RE (2011) Evolution of photosynthesis. *Annu Rev Plant Biol* 62:515–548.
2. Rae BD, Long BM, Badger MR, Price GD (2013) Functions, compositions, and evolution of the two types of carboxysomes: Polyhedral microcompartments that facilitate CO_2 fixation in cyanobacteria and some proteobacteria. *Microbiol Mol Biol Rev* 77:357–379.
3. Burnap RL, Hagemann M, Kaplan A (2015) Regulation of CO_2 concentrating mechanism in cyanobacteria. *Life (Basel)* 5:348–371.
4. Shibata M, et al. (2002) Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: Function and phylogenetic analysis. *J Biol Chem* 277:18658–18664.
5. Wang HL, Postier BL, Burnap RL (2004) Alterations in global patterns of gene expression in *Synechocystis* sp. PCC 6803 in response to inorganic carbon limitation and the inactivation of ndhR, a LysR family regulator. *J Biol Chem* 279:5739–5751.
6. Schwarz D, Schubert H, Georg J, Hess WR, Hagemann M (2013) The gene sml0013 of *Synechocystis* species strain PCC 6803 encodes for a novel subunit of the NAD(P)H oxidoreductase or complex I that is ubiquitously distributed among cyanobacteria. *Plant Physiol* 163:1191–1202.
7. Du J, Förster B, Rourke L, Howitt SM, Price GD (2014) Characterisation of cyanobacterial bicarbonate transporters in *E. coli* shows that SbtA homologs are functional in this heterologous expression system. *PLoS One* 9:e115905.
8. Forchhammer K, Lüddecke J (2016) Sensory properties of the P_{II} signalling protein family. *FEBS J* 283:425–437.
9. Fokina O, Chellamuthu VR, Zeth K, Forchhammer K (2010) A novel signal transduction protein P(II) variant from *Synechococcus elongatus* PCC 7942 indicates a two-step process for NAGK-P(II) complex formation. *J Mol Biol* 399:410–421.
10. Lapina T, Selim KA, Forchhammer K, Ermilova E (2018) The P_{II} signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida P_{II} proteins. *Sci Rep* 8:790.
11. Wheatley NM, et al. (2016) A PII-like protein regulated by bicarbonate: Structural and biochemical studies of the carboxysome-associated CP_{II} protein. *J Mol Biol* 428:4013–4030.
12. Terauchi K, Ohmori M (1999) An adenylate cyclase, Cya1, regulates cell motility in the cyanobacterium *Synechocystis* sp. PCC 6803. *Plant Cell Physiol* 40:248–251.
13. Huergo LF, Chandra G, Merrick M (2013) P(II) signal transduction proteins: Nitrogen regulation and beyond. *FEMS Microbiol Rev* 37:251–283.
14. Shibata M, et al. (2001) Distinct constitutive and low-CO_2-induced CO_2 uptake systems in cyanobacteria: Genes involved and their phylogenetic relationship with homologous genes in other organisms. *Proc Natl Acad Sci USA* 98:11789–11794.
15. Price GD (2011) Inorganic carbon transporters of the cyanobacterial CO_2 concentrating mechanism. *Photosynth Res* 109:47–57.
16. Price GD, Woodger FJ, Badger MR, Howitt SM, Tucker L (2004) Identification of a SulP-type bicarbonate transporter in marine cyanobacteria. *Proc Natl Acad Sci USA* 101:18228–18233.
17. Lüddecke J, Forchhammer K (2015) Energy sensing versus 2-oxoglutarate dependent ATPase switch in the control of *Synechococcus* P_{II} interaction with its targets NAGK and PipX. *PLoS One* 10:e0137114.

18. Fokina O, Chellamuthu VR, Forchhammer K, Zeth K (2010) Mechanism of 2-oxoglutarate signaling by the *Synechococcus elongatus* P_{II} signal transduction protein. *Proc Natl Acad Sci USA* 107:19760–19765.
19. Radchenko MV, Thornton J, Merrick M (2013) P(II) signal transduction proteins are ATPases whose activity is regulated by 2-oxoglutarate. *Proc Natl Acad Sci USA* 110: 12948–12953.
20. Michelet L, et al. (2013) Redox regulation of the Calvin-Benson cycle: Something old, something new. *Front Plant Sci* 4:470.
21. Tamoi M, Shigeoka S (2015) Diversity of regulatory mechanisms of photosynthetic carbon metabolism in plants and algae. *Biosci Biotechnol Biochem* 79:870–876.
22. Chen Y, et al. (2000) Soluble adenylyl cyclase as an evolutionarily conserved bicarbonate sensor. *Science* 289:625–628.
23. Cann MJ, Hammer A, Zhou J, Kanacher T (2003) A defined subset of adenylyl cyclases is regulated by bicarbonate ion. *J Biol Chem* 278:35033–35038.
24. Steegborn C, Litvin TN, Levin LR, Buck J, Wu H (2005) Bicarbonate activation of adenylyl cyclase via promotion of catalytic active site closure and metal recruitment. *Nat Struct Mol Biol* 12:32–37.
25. Hammer A, Hodgson DR, Cann MJ (2006) Regulation of prokaryotic adenylyl cyclases by CO_2. *Biochem J* 396:215–218.
26. Hennon GMM, et al. (2015) Diatom acclimation to elevated CO_2 via cAMP signalling and coordinated gene expression. *Nat Clim Chang* 5:761–765.
27. Hood RD, Higgins SA, Flamholz A, Nichols RJ, Savage DF (2016) The stringent response regulates adaptation to darkness in the cyanobacterium *Synechococcus elongatus*. *Proc Natl Acad Sci USA* 113:E4867–E4876.
28. Puszynska AM, O'Shea EK (2017) ppGpp controls global gene expression in light and in darkness in *S. elongatus*. *Cell Rep* 21:3155–3165.
29. Schwarz D, et al. (2011) Metabolic and transcriptomic phenotyping of inorganic carbon acclimation in the cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant Physiol* 155:1640–1655.
30. Heinrich A, et al. (2006) Interaction of the membrane-bound GlnK-AmtB complex with the master regulator of nitrogen metabolism TnrA in *Bacillus subtilis*. *J Biol Chem* 281:34909–34917.
31. Radchenko MV, Thornton J, Merrick M (2010) Control of AmtB-GlnK complex formation by intracellular levels of ATP, ADP, and 2-oxoglutarate. *J Biol Chem* 285: 31037–31045.
32. Huergo LF, et al. (2006) Interactions between P_{II} proteins and the nitrogenase regulatory enzymes DraT and DraG in *Azospirillum brasilense*. *FEBS Lett* 580:5232–5236.
33. Sassone-Corsi P (2012) The cyclic AMP pathway. *Cold Spring Harb Perspect Biol* 4: a011148.
34. Ohmori M, Okamoto S (2004) Photoresponsive cAMP signal transduction in cyanobacteria. *Photochem Photobiol Sci* 3:503–511.
35. Forcada-Nadal A, Forchhammer K, Rubio V (2014) SPR analysis of promoter binding of *Synechocystis* PCC6803 transcription factors NtcA and CRP suggests cross-talk and sheds light on regulation by effector molecules. *FEBS Lett* 588:2270–2276.
36. Nambi S, Basu N, Visweswariah SS (2010) cAMP-regulated protein lysine acetylases in mycobacteria. *J Biol Chem* 285:24313–24323.
37. Vagin A, Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* 66:22–25.
38. Orf I, et al. (2015) Integrated analysis of engineered carbon limitation in a quadruple CO_2/HCO_3^- uptake mutant of *Synechocystis* sp. PCC 6803. *Plant Physiol* 169:1787–1806.
39. Pinto FL, Thapper A, Sontheim W, Lindblad P (2009) Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Mol Biol* 10:79.
40. Imashimizu M, Yoshimura H, Katoh H, Ehira S, Ohmori M (2005) NaCl enhances cellular cAMP and upregulates genes related to heterocyst development in the cyanobacterium, *Anabaena* sp. strain PCC 7120. *FEMS Microbiol Lett* 252:97–103.

The FEBS Journal
FEBS PRESS
science publishing by scientists

Interaction of N-acetyl-L-glutamate kinase with the PII signal transducer in the non-photosynthetic alga Polytomella parva: Co-evolution towards a hetero-oligomeric enzyme

Journal:	*The FEBS Journal*
Manuscript ID	FJ-19-0164
Manuscript Type:	Regular Paper
Date Submitted by the Author:	12-Feb-2019
Complete List of Authors:	Selim, Khaled; University of Tuebingen, Inst. fur Mikrobiologie Lapina, Tatjana; Biological Faculty, Saint-Petersburg State University Forchhammer, Karl; University of Tuebingen, Inst. fur Mikrobiologie Ermilova, Elena; Biological Faculty, Saint-Petersburg State University
Key Words:	

Interaction of N-acetyl-L-glutamate kinase with the PII signal transducer in the non-photosynthetic alga *Polytomella parva*: Co-evolution towards a hetero-oligomeric enzyme

Khaled A. Selim[1], Tatyana Lapina[2], Karl Forchhammer[1] and Elena Ermilova[2]

[1]Organismic Interactions Department, Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany.

[2]Biological Faculty, Saint-Petersburg State University, Universitetskaya nab. 7/9, Saint-Petersburg 199034, Russia.

Authors For correspondence:

Karl Forchhammer (karl.forchhammer@uni-tuebingen.de),

Telefon +49 7071 29-72096
Telefax +49 7071 29 - 5843

Elena Ermilova (e.ermilova@spbu.ru)

Tel.: 007 812 4506740

fax: 007 812 4507310.

Authors email list:

Khaled A. Selim: Khaled.selim@uni-tuebingen.de;

Tatyana Lapina: t.lapina@spbu.ru;

Karl Forchhammer: karl.forchhammer@uni-tuebingen.de;

Elena Ermilova: e.ermilova@spbu.ru

Abstract (248/250)

During evolution, several algae and plants became heterotrophic and lost photosynthesis, however, in most cases, a non-photosynthetic plastid was maintained. Among these organisms, the colorless alga *Polytomella parva* is a special case, as its plastid is devoid of any DNA, but is maintained for specific metabolic tasks carried out by nuclear-encoded enzymes. This makes *P. parva* attractive to study molecular events underlying the transition from autotrophic to heterotrophic lifestyle. Here we characterize metabolic adaptation strategies of *P. parva* in comparison to the closely-related photosynthetic alga *Chlamydomonas reinhardtii* with a focus on the role of chloroplast-localized PII signaling protein. *P. parva* accumulates significantly higher amounts of most TCA cycle intermediates as well as glutamate, aspartate and arginine, the latter being specific for the colorless plastid. Correlating with the altered metabolite status, the carbon/nitrogen sensory PII signaling protein and its regulatory target N-acetyl-L-glutamate-kinase (NAGK; the controlling enzyme of arginine biosynthesis) show unique features: They have co-evolved into a stable hetero-oligomeric complex, irrespective of effector molecules. The PII signaling protein, so far known as a transiently interacting signaling protein, appears as a permanent subunit of the enzyme NAGK. NAGK requires PII to properly sense the feedback inhibitor arginine, and moreover, PII tunes arginine-inhibition in response to glutamine. No other PII effector molecules interfere, indicating that the PII-NAGK system in *P. parva* has lost the ability to estimate the cellular energy and carbon status but has specialized to provide an entirely glutamine-dependent arginine-feedback control, highlighting the evolutionary plasticity of PII signaling system.

Key-words:

Arginine biosynthesis,
algal metabolomics,
N-acetyl-L-glutamate kinase,
Nonphotosynthetic plastids,
PII-signaling,
TCA/GS-GOGAT cycles.

Introduction

The loss of photosynthesis is always accompanied with heterotrophic lifestyles and arose in diverse eukaryotic lineages [1]. In the course of evolution, many algal species and land plants lost photosynthesis and became heterotrophic [1-5]. Most of these non-photosynthetic organisms still retain the plastids, which contain a small genome to carry out various non-photosynthetic metabolic reactions [1,8]. Several of these colorless algae evolved into parasites, such as the Apicomplexa or the Oomycete lineage. *Polytomella* is a genus of colorless, free-living unicellular non-photosynthetic green algae, closely related to the photosynthetic green alga *Chlamydomonas reinhardtii* [6,7]. Recently, the plastids of *Polytomella* spp. have been identified to be the first example of plastid-harboring algae devoid of any plastid genomes [1].

RNA-seq analysis of *Polytomella parva* uncovered transcripts for a large set of nuclear encoded, plasmid-targeted enzymes mainly involved in carbohydrate and starch metabolism as well as amino acid and fatty acid biosynthesis [1]. This implies that *P. parva* has maintained a non-photosynthetic plastid for metabolic purposes as a specialized anabolic organelle [1,2,6,7]. Therefore, *P. parva* is an attractive model system for exploring the evolutionary pressure to maintain plastids in the absence photosynthesis. Up to now, the question how the primary metabolism in *Polytomella* spp., has adapted to the loss of photosynthesis has not been experimentally approached. Therefore, we started a first characterization concerning the biochemical and metabolic adaptation strategies of *P. parva* in response to different nitrogen regimes. Notably, *P. parva* was found to possess nuclear genes predicted to encode a plastid-targeted PII signaling protein (*Ppa*PII, plastid-targeted) and the enzymes of the ornithine/arginine biosynthesis pathway, in particular the target of PII regulation, N-acetyl-L-glutamate kinase (EC 2.7.2.8) (*Ppa*NAGK, plastid-targeted), which catalyzes the commented step of arginine biosynthesis. These two proteins co-evolved in the course of endosymbiotic generation of chloroplasts [9] and therefore, represent a prominent test case to address issues of metabolic adaptation strategies.

The PII signalling proteins constitute a large superfamily occurring in all domains of life [10,11]. The PII proteins are trimeric in the structure and are present in almost all bacteria, in nitrogen-fixing archaea [10,11] and in the eukaryotic Archaeplastida domain [9-12]. The PII homologues (GlnB and GlnK), which contain the conserved PROSITE motifs, are referred as canonical PII proteins. The PII members, which demonstrate the same trimeric architectural principle as GlnB/GlnK proteins but lack their typical PROSITE signature pattern, are termed as the PII-like proteins [10,13].

In contrast to the high structural conservation of PII proteins, the PII controlled targets are distinct and versatile in different phylogenetic lineages. In eukaryotes, PII homologues have only been identified and characterized in Chloroplastida (green algae and land plants), where they are nuclear-encoded [14-16] and in Rhodophyta, where they are coded by the plastid genome [12,17]. In both groups of eukaryotic phototrophs, PII is localized in the chloroplast [14-17]. In cyanobacteria and plants, the PII signalling proteins was found to regulate the activity of NAGK, the controlling enzyme of arginine biosynthesis [16-19]. In green algae and land plants, NAGK activity is controlled by the cellular glutamine (Gln) levels via glutamine-dependent PII-NAGK complex formation, which leads to increased enzyme activity [14,16]. In contrast to PII proteins from Chloroplastida, PII of the red alga *Porphyra purpurea* controls NAGK in a similar way as shown in cyanobacteria: PII-NAGK complex formation is antagonized 2-oxoglutarate (2-OG) but independent of glutamine. Through complex formation with PII, NAGK gets relieved from feedback inhibition by arginine (Arg) [17], leading to enhanced activity. It appears that the biochemical features of PII-NAGK complexes reflect the metabolic adaptations during endosymbiotic evolution.

The present study is the first to address metabolic adaptation strategies of the nonphotosynthetic alga *P. parva* in response to nitrogen limitation in comparison to the closely related photosynthetic alga *C. reinhardtii* by performing a relative quantification of the intracellular metabolites. To gain mechanistic insights in the metabolic specialization of the *P. parva* plastid, we studied the PII-mediated regulation of NAGK activity, which is a key step in the control of arginine biosynthesis. Surprisingly, we found unique features not described for PII-NAGK complexes so far. *Ppa*PII forms an unusually stable complex with *Ppa*NAGK, irrespective of effector molecules. In this complex, PII tunes arginine feedback-inhibition of NAGK specifically in response to varying glutamine levels, whereas the tricarboxylic acid (TCA) cycle intermediate 2-oxoglutarate (2-OG) and ATP/ADP nucleotides had no regulatory effect. These data indicate that the PII-NAGK system in this non-photosynthetic alga evolved into a hetero-oligomeric enzyme complex that has lost the ability to estimate the current energy and carbon status of the cells but specifically responds with high sensitivity to the arginine/glutamine level of the cells.

Results

Metabolomic analysis.

To investigate the impact of the non-photosynthetic lifestyle of *P. parva* on its metabolomic landscape, we applied an untargeted LC-MS metabolomics approach to characterize the changes in the metabolomic pool sizes of *P. parva* cells under different nitrogen regime (Fig. S1) in

comparison to the closely related alga *C. reinhardtii* grown under optimal photosynthetic conditions. We were able to identify 11 metabolites of the central C and N metabolism, mainly of TCA and GS-GOGAT cycles (Fig. 1 and Table S1), which were significantly different between *C. reinhardtii* and *P. parva* and changed upon shift from high to low nitrogen.

Remarkably, the pools of most tricarboxylic acid (TCA) cycle intermediates (citrate, isocitrate, succinate and fumarate), except malate and 2-OG were much higher in *P. parva* cells (Fig. 1). The high levels of TCA intermediates in *P. parva* cells can be explained by the organotrophic life-style of *P. parva*, which uses ethanol as a carbon source, oxidatively metabolizing it by mitochondrial activity, as an adaptation to the heterotrophic life-style (i.e. *P. parva* is mainly dependent on mitochondrial reactions for energy release). In striking contrast to most TCA intermediates, the levels of 2-OG were lower in *P. parva* than in *C. reinhardtii*. This suggests an efficient nitrogen assimilatory system in *P. parva* that constantly keeps the 2-OG levels relatively low, as compared to other TCA cycle intermediates. During nitrogen deprivation, the 2-OG level increases in *P. parva*, as expected [20], since the consumption of 2-OG through nitrogen assimilatory reactions is reduced. Intriguingly, the levels of phosphoenolpyruvate (PEP) show the inverse pattern than most TCA intermediates. The low level of PEP in *P. parva* (7.6-fold decrease as compared to *C. reinhardtii*) may be associated with the loss of photosynthetic metabolism, since PEP is synthesized in photoautotrophic algae from the CO_2 fixation product 3-phosphoglycerate (3PGA) through a few glycolytic reactions [21]. By contrast, in the heterotrophic *P. parva*, PEP has to be synthesized via gluconeogenetic reactions starting from the carbon source ethanol.

Under nitrogen rich conditions, *P. parva* cells accumulate around 2.9-fold more arginine, 9.4-fold more aspartate and 3.1-fold more glutamate than the *C. reinhardtii* cells. This suggests again, in agreement with lower levels of 2-OG, an efficient nitrogen assimilatory system. Nitrogen assimilation and arginine synthesis takes place in the colorless plastid, as the corresponding enzymes glutamine synthase (GS) and glutamate synthase (GOGAT) as well as arginine biosynthesis enzymes are plastid localized. In contrast to the elevated levels of Glu, Asp and Arg in *P. parva*, the Gln-levels were relatively low, which suggests a high activity of GOGAT, that constantly consumes glutamine and 2-OG to produce glutamate. The high levels of Glu correlate with high Arg levels, indicating that the controlling enzyme of the ornithine/arginine pathway, *Ppa*NAGK, should be adapted to the specific metabolic alterations in *P. parva*.

Upon shift of *P. parva* cells from nitrogen-rich to poor conditions, marked changes were mainly observed for metabolites of the TCA and GS/GOGAT cycles. The arginine, succinate and fumarate pools dropped by more than 50% (Fig. 1), whereas the malate, glutamate, and

glutamine (the primary nitrogen assimilation product) pools dropped by 30% to 40% (Fig. 1). The amount of aspartate increased slightly, which can be explained by diminished aspartate consumption for arginine synthesis through the argininosuccinate synthase reaction. As expected, the central TCA product 2-OG showed a more than 2-fold increase upon shift to low nitrogen condition, whereas the 2-OG precursors citrate and isocitrate did not show marked changes. The three-fold increase of PEP levels under N-limitation reflects the shift in the C:N ratio during external N-limitation. Due to the limitation of nitrogen assimilation reactions under N-poor conditions, the decreased utilization of glycolytic intermediates for various amino acid biosynthesis reactions could lead to increased levels of the glycolytic metabolite PEP. Overall, these metabolic changes reflect the limitation of nitrogen availability, which goes along with a slightly reduced growth of *P. parva* cells under these nitrogen-poor conditions (Fig. S1).

Together, the main metabolic difference between the photosynthetic alga *C. reinhardtii* and its heterotrophic relative *P. parva* concerns major metabolites of the TCA cycle, and nitrogen assimilation products glutamate, aspartate as well as arginine as a final N-storage molecule [22]. The higher levels of TCA intermediates agree with the dominance of mitochondrial metabolism in *P. parva*. The high levels of nitrogen assimilation products, in particular the nitrogen-storage amino acid arginine, which is produced in the color-less plastid, indicates a prominent metabolic role of this organelle. Since the master nitrogen regulatory PII protein is crucial for arginine synthesis and senses the cellular C:N balance [10], we hypothesized that the *P. parva* PII orthologue (*Ppa*PII) could play a crucial role for its metabolic adaptation strategy. To examine this hypothesis, we decided to study the *P. parva* PII-mediated regulation of NAGK system, regarding the activation of the committed step of arginine biosynthesis.

_Ppa_PII is a canonical PII protein.

The predicted full-length *Ppa*PII polypeptide encoded by the *P. parva GLB1* gene consists of 209 amino acids with a calculated molecular weight of 22745 Da and contains predicted chloroplast transit peptide using ChloroP 1.1 Server (amino acid residues 1-49). As expected, the mature *Ppa*PII demonstrated the highest degree of identity with *C. reinhardtii* PII (61.78%). We performed primary sequence alignment of PII from *P. parva* with canonical PII proteins from other Archaeplastida and bacteria. The alignment of *Ppa*PII indicates extremely high local identities over two signature patterns that have been defined at the PROSITE (PS00496 and PS00638) in all canonical PII proteins (Fig. 2). Moreover, similar to PII homologues of Chloroplastida, *Ppa*PII protein contains the unique C-terminal segment, which is responsible for glutamine sensing [14,16]. The alignment also showed the presence of the signature pattern I (Y-[KR]-G-[AS]-[AE]-Y) that contains the residues involved in the interaction with NAGK. Indeed,

the *NAG1* gene was identified in the *P. parva* genome, encoding the full-length NAGK polypeptide (*Ppa*NAGK) that consists of 329 amino acids with a calculated molecular weight of 34819 Da with a putative N-terminal plastid transit peptide. The calculated molecular weight of the predicted mature *Ppa*NAGK polypeptide is 32587 Da.

To gain further insights about biochemical properties of *Ppa*PII and *Ppa*NAGK proteins, we prepared respective recombinant proteins. Therefore, a recombinant N-terminal His-tagged variant of the predicted mature *Ppa*NAGK protein without the chloroplast signal peptide (amino acid residues 1-40) (the theoretical molecular mass of monomeric recombinant *Ppa*NAGK protein is 32.6 kDa), and a recombinant Strep-tagged version of the mature *Ppa*PII protein without the chloroplast signal peptide (the theoretical molecular mass of monomeric recombinant *Ppa*PII protein is 19.4 kDa) were overexpressed in *E. coli* and affinity-purified.

_Ppa_NAGK catalytic efficiency in the absence of arginine is not influenced by PII.

The kinetic constants of the purified recombinant *Ppa*NAGK enzyme in the absence of the feedback inhibitor arginine were determined (Fig. 3A) and showed a K_m value for NAG of 2.35 ± 0.22 mM and a maximum K_{cat} of 211.5 ± 4.1 s^{-1}. In the presence of *Ppa*PII, the apparent K_m for NAG and K_{cat} slightly increased with values to 3.99 ± 0.62 mM and 385.4 ± 16.2 s^{-1}. Therefore, the overall catalytic efficiency (K_{cat}/K_m) was of 90 x 10^3 and 96.5 x 10^3 $s^{-1}M^{-1}$ for free and *Ppa*PII-complexed *Ppa*NAGK, respectively. Strikingly, addition of Gln did not cause any increase in K_{cat}/K_m (96.9 × 10^3 s^{-1} M^{-1}), in stark contrast to the situation in *C. reinhardtii* [16]. Together, the overall *Ppa*NAGK catalytic efficiency was not affected by *Ppa*PII, neither in the presence nor in the absence of Gln.

Arginine inhibition of _Ppa_NAGK activity and its modulation by glutamine.

As the relief from arginine inhibition by PII-NAGK complex formation is a checkpoint for metabolic control of arginine biosynthesis in cyanobacteria and Chloroplastida [16-19], we asked if the presence of *Ppa*PII could change the arginine inhibition profile of *Ppa*NAGK. In the absence of *Ppa*PII, feedback inhibition by arginine of *Ppa*NAGK occurred with a half maximal inhibitory concentration (IC_{50}) of 0.67 ± 0.04 mM (Fig. 4A). Interestingly, addition of *Ppa*PII protein to *Ppa*NAGK enhanced the arginine sensitivity of NAGK by dropping the IC_{50} for arginine by 3.7-fold to 0.18 ± 0.01 mM. Addition of glutamine to the assay mixture in presence of *Ppa*PII led to increased *Ppa*NAGK activity at low concentrations of Arg and strongly relieved the complex from Arg feedback-inhibition. By contrast, in the absence of *Ppa*PII, glutamine had no influence on NAGK activity (Fig. 4A).

To investigate further whether the increased sensitivity of the *Ppa*NAGK-PII complex towards Arg, as compared to *Ppa*NAGK alone, is a property of *Ppa*PII or of *Ppa*NAGK, we performed heterologous enzymatic assays using the respective *C. reinhardtii* proteins. Furthermore, we tested the PII protein from rice plant *Oryza sativa* (*Os*PII) (*Cr*PII) over *Ppa*NAGK. The *Cr*NAGK protein is inherently sensitive towards arginine. Addition of *Ppa*PII protein to *Cr*NAGK further increased the sensitivity of *Cr*NAGK for arginine inhibition, by dropping the IC_{50} for arginine from 0.12 ± 0.03 to 0.07 ± 0.01 in the absence and presence of *Ppa*PII, respectively (Fig. 4B). In the presence of 5 mM glutamine, *Ppa*PII behaved as shown previously for *Cr*PII (Fig. S2) and [16], enhancing the enzymatic activity of *Cr*NAGK and strongly relieving arginine feedback inhibition, as evidenced by the 4-fold increase of the IC_{50} for arginine to 0.48 ± 0.04 mM (Fig. 4B). Conversely, the *Cr*PII and *Os*PII proteins did not raise arginine sensitivity of the *Ppa*NAGK (Fig. 4C,D). In the absence of Gln, the *Cr*PII and *Os*PII proteins slightly enhanced *Ppa*NAGK activity at low concentrations of Arg (0 to 0.5 mM). At high arginine concentrations (in the absence of Gln), *Ppa*NAGK activity dropped in presence or absence of the heterologous PII proteins (corresponding to IC_{50} values of 0.68 ± 0.15 and 1.5 mM, respectively) (Fig 4C,D). In the presence of 5 mM glutamine, *Cr*PII and *Os*PII proteins relieved arginine feedback inhibition of *Ppa*NAGK, as expected [16] with IC_{50} values of 1.2 and 2.3 mM, respectively. These results showed that the *Ppa*PII-mediated sensitivation of *Ppa*NAGK activity towards arginine is a unique property of *Ppa*PII.

Because glutamine increases the activity of the *Ppa*PII-NAGK complex (Fig. 4), we further tested the activation of arginine-inhibited *Ppa*NAGK-*Ppa*PII complex by glutamine in a concentration-dependent manner (Fig. 5A) in presence of 0.5 mM Arg. The half-maximal effective concentration (EC_{50}) of glutamine for activation of the *Ppa*PII-NAGK complex was determined as 1.8 mM. Strikingly, this response was very similar to the glutamine-dependent activation of *Cr*NAGK by *Ppa*PII, with a glutamine EC_{50} of 1.1 mM (Fig. 5A). By comparison, the EC_{50} of glutamine for stimulation of *Cr*NAGK activity by *Cr*PII was 2.4 ± 0.8 mM [16]. This shows that *Ppa*PII has evolved to sense lower glutamine concentrations.

_Ppa_PII protein lacks the response to 2-oxoglutarate.

Most PII proteins were found to sense 2-OG as the principle effector molecule in a synergistic binding reaction with ATP [10,20,23]. Recent studies have identified PII proteins from some Chloroplastida that lack 2-OG responses [17]. We next assessed the significance of 2-OG for interaction of *Ppa*PII and *Ppa*NAGK. As shown in Fig. (5B), addition of 2-OG to the *Ppa*PII-*Ppa*NAGK-Arg-Gln mixture did not lead to inhibition of *Ppa*NAGK activity, which would be expected if the complex would dissociate. As a control, the expected response toward 2-OG was

obtained for the heterologous assay with *Cr*PII, which senses 2-OG [16,17] with an IC_{50} of 1.99 mM. Unlike *Cr*PII, the higher plant *Os*PII showed a very weak response towards 2-OG. Together, it appears that the *Ppa*PII protein does not sense 2-OG, unlike *Cr*PII [16,17] but complexes NAGK to tune its response towards the feedback-inhibitor arginine in a glutamine-dependent manner.

Glutamine-independent *Ppa*PII-*Ppa*NAGK complex formation.

The above described enzyme tests suggested that *Ppa*PII-NAGK complex formation must be different from all previously tested cases [16-19,22], since *Ppa*PII enhances the arginine sensitivity of NAGK in the absence of glutamine. Of note, in the absence of Gln, the PII proteins from representative Chloroplastida were not able to effectively form a complex with NAGK, even in the presence Mg^{2+}-ATP [14,16]. *Ppa*PII contains the C-terminal Q-loop responsible for glutamine binding (Fig. 2), that was shown previously to promote glutamine-dependent complex formation of *Cr*PII-NAGK. To monitor any small molecular weight shift due to the complex formation, we tried to characterize PII-NAGK complexes in presence or absence of Gln using analytical size exclusion chromatography (SEC) coupled to multi angle light scattering (MALS). First, we determined the oligomerization state of *Ppa*PII and *Ppa*NAGK proteins. As expected, the *Ppa*PII protein eluted as a trimer and *Ppa*NAGK as a hexamer (Fig. 3B). When an excess of *Ppa*PII was mixed with *Ppa*NAGK (4:1 monomeric concentrations), a PII-NAGK complex was detected with a clearly detectable peak shift for the *Ppa*NAGK hexamer. Of note, glutamine was not required for complex formation, nor did it induce a remarkable shift in the size of the complex (Fig. 3B).

To find out whether it is a property of the *Ppa*PII protein, which is responsible for glutamine-independent complex formation with NAGK, we investigated complex formation with the *Cr*NAGK protein. The *Cr*NAGK eluted as a hexamer like previously reported [16]. Independent of the absence or presence of Gln, the *Ppa*PII protein was able to form a stable complex with *Cr*NAGK and both proteins eluted together (Fig. S3). In agreement, the SDS-PAGE analysis of the collected peaks showed the presence of both *Ppa*PII with different NAGKs in the complex peak (Fig. 3C, S3). Together, the results demonstrated that the *Ppa*PII protein forms complexes with NAGKs independent of glutamine.

Influence of different effector molecules on PII-NAGK complex formation.

To confirm that the direct interaction between *Ppa*PII and *Ppa*NAGK is glutamine-independent and to test the influence of the other known PII effectors molecules (ATP, ADP, 2-OG, and Gln) on the PII-NAGK complex formation, we assessed the complex formation using surface plasmon

resonance (SPR) spectroscopy. In SPR experiments, the His-tagged NAGK protein was immobilized on a Ni-NTA sensor chip and the strep-tagged PII protein was injected together with or without different combinations of effectors molecules to monitor the difference in the response unites (ΔRU) due to the PII-NAGK complex formation.

We showed previously that formation of the PII-NAGK complex from *C. reinhardtii* was strictly Mg^{2+}-ATP and glutamine dependent and was not supported by ADP (Fig. S4 and [16]). By contrast, *Ppa*PII was able to form a strong complex with NAGK on the SPR surface, independent of presence or absence of ADP, ATP and Gln (Fig. 6A). Remarkably, the *Ppa*NAGK-*Ppa*PII complex was extraordinary stable and dissociated very slowly in the course of the assay with estimated K_d value of 93.8 ± 29.9 nM (Fig. S4A,B). However, in the case of *Cr*PII-*Cr*NAGK, the complex dissociated spontaneously at the end of the injection as soon as it encountered a buffer devoid of Mg^{2+}-ATP and Gln (compare Fig. S4C with S4A). Intriguingly, the percent of *Ppa*PII-NAGK complex dissociation from the sensor chip at 330 and 660 sec after the end of the injection phase was 60.7% and 41.9%, respectively (RU at 110 sec was taken as 100%) (Fig. S4B), indicating the stability of the complex.

Furthermore, we reported previously that ADP and 2-OG negatively affected cyanobacterial PII-NAGK interaction by promoting the dissociation of the complex, and further, the injection of 1mM of ADP caused immediate dissociation of PII–NAGK complexes [24]. Unexpectedly, the *Ppa*NAGK-PII complex was resistant against the injection of ADP, hinting that *Ppa*PII may be unable to sense ADP (Fig. 6A, S4A). The 2-OG effector also showed no antagonistic effect on *Ppa*PII-NAGK complex formation (Fig. 6A), in agreement with the inability of 2-OG to inhibit the NAGK enzymatic activity (Fig. 5B). This result resembled previous results on *Cr*PII-*Cr*NAGK interaction, where 2-OG had no influence on the NAGK-PII complex formation, while the 2-OG mediated inhibition of *Cr*NAGK activity in complex with *Cr*PII-complex appeared to occur post-binding [16].

To confirm that the ability of *Ppa*PII to form a Gln-independent complex with NAGK is an intrinsic property of this PII protein, we repeated the previous SPR experiments using *Cr*NAGK. Regardless of the effector molecules added to the assay mixture (containing 1 μM *Ppa*PII protein), the *Ppa*PII protein formed a strong complex with *Cr*NAGK in a Mg^{2+}-ATP/Gln independent manner, and similarly ADP and 2-OG had no influence on *Ppa*PII-*Cr*NAGK complex formation (Fig. 6B). Again, *Ppa*PII was able to bind to *Cr*NAGK without any effector molecules (K_d value of 86.3 ± 9.4 nM, Fig. 6C). From the previous experiments, we concluded the Gln independency of PII-NAGK complex formation, are unique features of *Ppa*PII.

We then asked whether the *Ppa*NAGK protein provides features to the glutamine-independence of *Ppa*PII-NAGK complex formation. Therefore, we tested the ability of *Cr*PII and *Os*PII

proteins to form heterologous complexes with *Ppa*NAGK in presence or absence of effector molecules Mg^{2+}-ATP and Gln using SPR. As already mentioned, the formation of the *Cr*NAGK-*Cr*PII complex is strictly dependent on Mg^{2+}-ATP and Gln (Fig. S4C), whereas Mg-ADP did not support complex formation [16]. Remarkably, *Cr*PII and *Os*PII proteins were able to bind to *Ppa*NAGK without any effector molecules (Fig. 7A, S4D), indicating that *Ppa*NAGK attracts the heterologous PII proteins in a glutamine independent manner. Nevertheless, with *Cr*PII, the presence of Mg-ATP alone or in combination Gln moderately or strongly enhanced the binding to *Ppa*NAGK, respectively. Moreover, *Os*PII interacted with *Ppa*NAGK independent of any effector molecules (Fig. S4D). Interestingly, in the presence of Mg-ADP, *Cr*PII was still able to form a weak complex with *Ppa*NAGK, similar to the absence of effector molecules, indicating that *Cr*PII lost the ability to sense ADP, confirming our previous reports (Fig. 7B) [16,17]. The addition of Gln in presence of Mg-ADP enhanced the *Cr*PII-*Ppa*NAGK complex formation, however, Gln in combination with Mg-ATP simulated much stronger complex *Cr*PII-*Ppa*NAGK formation (compare Fig 7A and 7B). Again, 2-OG did not show any influence on the *Cr*PII-*Ppa*NAGK complex (Fig. 7C). These results indicate that *Ppa*NAGK strongly influences the binding properties of various PII proteins and implies a role in the sensory properties of the entire PII-NAGK complex.

Discussion

The *Polytomella* spp. lineages including *P. parva* contain a strongly reduced mitochondrial genome and more strikingly, are the first discovered plastid-bearing organisms devoid of a plastid genome [1]. Apparently, the *P. parva* plastid seems to carry out essential anabolic functions including amino acid, fatty acid, carbohydrate and lipid biosynthesis, which have not been re-located to the cytoplasm during evolution. In *P. parva*, the genes encoding for NAGK and PII proteins were found among the plastid-targeted/nuclear-encoded genes. We hypothesized that *P. parva* must possess strong metabolic adaptations to cope with the evolutionary loss of photosynthesis.

As a consequence of the life-style switch towards a purely heterotrophic metabolism, the mitochondria in *P. parva* are the primary energy-generating organelle through their respiratory activity, whereas in *C. reinhardtii*, their respiratory function is mainly limited to dark periods. In agreement with the prominent role of mitochondrial metabolism, the levels of most of the TCA cycle intermediates in *P. parva* are strongly increased as compared to *C. reinhardtii*. By contrast, in *C. reinhardtii,* the elevated PEP pool (as compared to *P. parva)* agrees with a flow of carbon from CO_2 fixation into lower glycolysis [25], to provide the cells with precursors for most anabolic pathways. Under conditions of nitrogen excess, the levels of the amino acids: arginine,

aspartate and glutamate are clearly elevated *P. parva* compared with *C. reinhardtii* (Fig. 1), which suggests fast nitrogen-assimilation reactions in *P. parva*. The GS/GOGAT cycle is the primary route for nitrogen assimilation, suggesting that this reaction cycle should be highly active in *P. parva*, in agreement with the localization of GS and GOGAT enzymes in the *P. parva* plastid and the fast growth rate of *P. parva* in comparison to *C. reinhardtii* [26,27]. As compared to *C. reinhardtii*, the level of 2-OG is relatively the lowest of all TCA cycle metabolites in *P parva*. This agrees with an efficient GOGAT reaction, constantly depleting the 2-OG pool. Since this reaction also consumes glutamine, we also find relatively lower levels of glutamine than of glutamate, aspartate or arginine. The highly active nitrogen assimilation activity results not only in elevated glutamate levels but also in high arginine levels. The controlling enzyme of the arginine synthesis pathway is NAGK, therefore, needs to be highly active. As shown here, the control of NAGK by the PII signalling protein shows unique features, which probably result from the evolutionary pressure of a non-photosynthetic environment with a restrained adenylate energy charge to maintain NAGK at high activity. These adaptations are discussed in the following:

Unlike all other PII-NAGK couples from cyanobacteria and plants investigated so far, the *Ppa*PII-NAGK complex associates in an almost irreversible manner. The effector molecules ADP or Mg-ATP/2-OG, which cause efficient dissociation of the complex, are ineffective in the case of the *P. parva* proteins. Even more strikingly, the complex forms in a completely glutamine-independent manner, although the glutamine-sensing C-terminal extension, the Q-loop [16] is perfectly conserved in *Ppa*PII (Fig. 2). Since the amino acid sequences of PII and NAGK are highly conserved between *P. parva* and *C. reinhardtii* (with 61.7% and 84.5% identity, respectively), these unique features were unexpected. A few amino acid substitutions may be sufficient to change the transient PII-NAGK complex, as it was characterized in all oxygenic phototrophs [16-19,22], into a stable hetero-oligomeric enzyme complex. In *P. parva,* the PII protein has turned into a stably attached regulatory subunit of NAGK. The very high *in vitro* stability of the complex suggests that these proteins are co-evolved towards the formation of a stable hetero-oligomeric complex in an effector molecule-independent manner, and probably are always complexed *in vivo*. In this complex, the Gln binding site resides in the PII subunit and NAGK exhibits the allosteric Arg site. However, the entire complex is required for sensitive reaction of NAGK towards Arg.

The regulatory effects of the effector molecules in the *P. parva* NAGK enzyme are apparently not mediated by complex formation or dissociation, but occur at the post-binding stage. At low Gln-levels (corresponding to N-poor conditions), the *Ppa*PII-NAGK complex is much more sensitive towards arginine than the isolated *Ppa*NAGK protein alone. This indicates that in

complex with *Ppa*PII, NAGK adopts a conformation that has high affinity for the allosteric feedback inhibitor arginine. By contrast, when Gln binds to the complex (under N-rich conditions), *Ppa*PII strongly relieves NAGK from arginine feedback-inhibition, indicating that glutamine imposes a conformational change on the entire complex that counteracts the feedback inhibitory effect of arginine, like previously shown for PII-NAGK complexes from other oxygenic phototrophs [16]. The half-maximal effective concentration (EC_{50}) of glutamine to stimulate *Ppa*PII-NAGK activity in presence of arginine is 50% lower than the EC_{50} for *Cr*PII, showing that the PII protein has evolved to allow enhanced NAGK activity at lower glutamine concentrations. Therefore, we speculate that the default mode of the PII-NAGK system in *P. parva* is a high arginine production through the extremely active *Ppa*NAGK-*Ppa*PII complex, unless there is a severe nitrogen (Gln) limitation.

Analysis of the heterologous complexes formed between *P. parva* proteins and plant PII or *Cr*NAGK proteins allowed us to conclude that in *P. parva*, both partner proteins have co-evolved towards a stable complex, with both proteins contributing to the enhanced complex stability. For this reason, *Cr*PII shows already measurable binding to *Ppa*NAGK even in the absence of glutamine, whereas with its proper partner *Cr*NAGK, binding is absolutely Mg-ATP and Gln-dependent. The high affinity of *Ppa*NAGK for PII proteins is also evident from the avid binding of *Cr*PII regardless of effector molecules that usually dissociate PII-NAGK complexes (Mg-ATP/2-OG or ADP). Conversely, *Ppa*PII is prone to bind to NAGK proteins irrespective of effector molecules, as demonstrated by the effector molecule-independent binding of *Ppa*PII to *Cr*NAGK (which usually only accepts glutamine-ligated *Cr*PII as partner). This shows that the *Ppa*PII protein has evolved to exclusively sense the glutamine level in a very sensitive manner. The loss of sensing the ATP/ADP ratio and of 2-OG might be attributed to the loss of photosynthetic activity in the plastid with consequent metabolic changes. Recently we found that the PII protein from the moss *Physcomitrella patens* [17] has also lost the ability to sense ADP and 2-OG. This suggests that the detailed sensing properties of the PII proteins can easily be adjusted to the regulatory need of the respective metabolic situation in an organism.

Collectively, our finding extends the knowledge of PII signaling in planta. Apparently, it seems that during the evolution of Viridiplantae, the PII proteins diverged in their properties, becoming very heterogeneous with respect to 2-OG and to ADP binding and toward complex formation with NAGK. *P. parva* is an extreme case, where the PII protein specialized its function towards a glutamine-regulated subunit of the key enzyme of the arginine pathway NAGK. Possibly, other targets of PII regulation might have been lost during the reductive evolution of the non-photosynthetic organelle, allowing PII to exclusively focus on NAGK regulation. It would be interesting in future to investigate PII-NAGK systems in other secondary non-photosynthetic

organisms, to reveal if the unique feature of the PII-NAGK complex in *P. parva* is related to the loss of photosynthesis during evolution.

Material and Methods

Strains and cultivation conditions.

The strains and the cultivation conditions are described in the online supporting information (Material and Methods S1).

Metabolite extraction and quantification.

For quantification the intracellular metabolites of 50 ml of exponentially growing cells under the day cycle of *P. parva* under different nitrogen regimes (excess nitrogen of 7.5 mM NH_4Cl or poor nitrogen of 0.375 mM NH_4Cl) and of *C. reinhardtii* under nitrogen rich condition (7.5 mM NH_4Cl) were harvested by centrifugation at 4°C and the cell pellets were immediately frozen in liquid nitrogen. The metabolites extraction and quantification were done according to [28]. Full metabolite extraction and quantification protocol is described in the online supporting information (Material and Methods S1).

Cloning of *Ppa*PII and *Ppa*NAGK-like proteins.

Gene Blocks, with optimized codon usage for cloning and expression into *E. coli*, encoding for amino acid sequences of mature *Ppa*NAGK-like and *Ppa*PII genes without chloroplast signal peptides, were synthesized by IDT, USA. The gene Blocks for *Ppa*NAGK and *Ppa*PII were cloned directly into NdeI-digested pET15b vector (Novagen, Germany) and BsaI-digested pASK-IBA3 vector (IBA, Germany), respectively. Primers and full cloning protocol are described in the online supporting information (Material and Methods S1).

Expression and purification of *Ppa*NAGK, *Cr*NAGK, *Ppa*PII, *Cr*PII and *Os*PII proteins.

The overexpression of the recombinant N-terminal fused His_6-tagged *Ppa*NAGK and *Cr*NAGK were performed in *E. coli* LEMO-21(DE3) and the proteins were affinity purified on a Ni-NTA columns according to [29]. Overexpression of the recombinant C-terminal fused Strep-tagged PII proteins (*Ppa*PII, *Cr*PII and *Os*PII) were performed in PII-deficient *E. coli* RB9060 [17] and the proteins were affinity purified on a Strep-Tactin II column according to [13,19].

Coupled NAGK activity assay.

The activity of NAGK was assessed using a coupled enzyme assay in which the production of ADP after the consumption of ATP for phosphorylation of NAG was associated with the

oxidation of NADH by pyruvate kinase and lactate dehydrogenase as described previously [17,18]. The enzymatic constants K_m, k_{cat}, IC_{50} and EC_{50} were calculated from the velocity slopes using the Graph-Pad Prism software program (GraphPad Software, USA). The detailed NAGK enzymatic assay protocol is described in the online supporting information (Material and Methods S1).

Surface plasmon resonance spectroscopy analysis (SPR spectroscopy).

SPR experiments were done at 25°C using a BIA-core-X biosensor system (Biacore AB, Uppsala, Sweden) in HBS buffer (10 mM HEPES, 150 mM NaCl, 2 mM $MgCl_2$ and 0.005% NP-40, pH 7.5) with a flowrate of 15µl/min, as described previously [16,29]. The recombinant His_6-tagged NAGKs (*Ppa*NAGK and *Cr*NAGK) proteins were immobilized on the flow cell (FC2) of the Ni^{2+}-loaded NTA-biosensor chip. Followed by injection of PII proteins (25 µl) into both FC1 (control, for unspecific binding of PIIs to the surface of the sensor chip) and FC2 (NAGKs immobilized) on the sensor chip. The detailed SPR protocol is described in the online supporting information (Material and Methods S1).

Size Exclusion Chromatography and Multiangle Light Scattering (SEC-MALS) Analysis.

Analytical size exclusion chromatography was carried out according to [30] on a micro-Äkta purifier system equipped with Superose 6 Increase 10/300 GL (GE Healthcare, column volume of 24 ml). The Superose column was coupled to a triple-angle light scattering (MALS) detector: miniDAWN™ TREOS® system (Wyatt Technology Corp., CA) and a refractometer (Optilab T-rEX, Wyatt). The running buffer consisted of 10 mM Tris pH 7.8, 300 mM NaCl, 2 mM $MgCl_2$, 0.02% NaN_3 and 2% glycerol. The samples were injected for analysis with a flow rate 0.5 ml/min. The resulting data were analyzed with ASTRA program (Wyatt Technology). The apparent molecular weights were derived from MALS data. The chromatographic elution profiles were collected (0.5 ml fractions) and analyzed by Glycine-SDS PAGE. The full SEC-MALS protocol is described in the online supporting information (Material and Methods S1).

Acknowledgments

This work was supported by Saint-Petersburg State University (Grant No. 1.65.38.2017) to E.E., and DFG to K.F. (Fo195/9-2, Fo195/13-1). The authors are grateful to Jörg Scholl and Heinz Grenzendorf (IMIT, Tübingen University) for the technical assistance, the staff of the analytics unite (ZMBP, Tübingen University), and Christine Kiefer (Schäffer lab, ZMBP, Tübingen University) for providing *C. reinhardtii* strain.

Author contributions

E.E. and K.F. conceived and initiated the project. K.A.S. and K.F. designed the experiments. K.A.S. and T.L. performed experiments. K.A.S. interpreted the results and wrote the first draft of the manuscript. K.A.S, E.E. and K.F. commented and edited on the manuscript. All authors analyzed the results and approved the final version of the manuscript.

References

1. Smith DR & Lee RW (2014) A plastid without a genome: evidence from the nonphotosynthetic green algal genus *Polytomella*. *Plant Physiol* **164**(4):1812-1819.
2. Asmail SR & Smith DR & (2016) Retention, erosion, and loss of the carotenoid biosynthetic pathway in the nonphotosynthetic green algal genus *Polytomella*. *New Phytol* **209**(3):899-903.
3. Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. *Philos Trans R Soc Lond B Biol Sci* **365**(1541):729-748.
4. Wicke S, Muller KF, dePamphilis CW, Quandt D, Wickett NJ, Zhang Y, Renner SS & Schneeweiss GM (2013) Mechanisms of functional and physical genome reduction in photosynthetic and nonphotosynthetic parasitic plants of the broomrape family. *Plant Cell* **25**(10):3711–3725.
5. Molina J, Hazzouri KM, Nickrent D, Geisler M, Meyer RS, Pentony MM, Flowers JM, Pelser P, Barcelona J, Inovejas SA *et al.* (2014) Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Mol Biol Evol* **31**(4):793–803.
6. Smith DR & Asmail SR (2014) Next-generation sequencing data suggest that certain nonphotosynthetic green plants have lost their plastid genomes. *New Phytol* **204**(1):7-11.
7. Smith DR, Hua J, Archibald JM & Lee RW (2013) Palindromic genes in the linear mitochondrial genome of the nonphotosynthetic green alga *Polytomella magna*. *Genome Biol Evol* **5**(9):1661-1667.
8. Pombert JF, Blouin NA, Lane C, Boucias D & Keeling PJ (2014) A lack of parasitic reduction in the obligate parasitic green alga *Helicosporidium*. *PLoS Genet* **10**(5):e1004355.
9. Chellamuthu VR, Alva V & Forchhammer K (2013) From cyanobacteria to plants: conservation of PII functions during plastid evolution. *Planta* **237**(2):451-462.
10. Forchhammer K & Lüddecke J (2016) Sensory properties of the PII signalling protein family. *FEBS J* **283**(3):425-437.
11. Huergo LF, Chandra G & Merrick M (2013) P(II) signal transduction proteins: nitrogen regulation and beyond. *FEMS Microbiol Rev* **37**(2):251-283.
12. Uhrig RG, Ng KK & Moorhead GB (2009) PII in higher plants: a modern role for an ancient protein. *Trends Plant Sci* **14**(9):505-11.
13. Selim KA, Haase F, Hartmann MD, Hagemann M & Forchhammer K (2018) PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. *Proc Natl Acad Sci U S A* **115**(21):E4861-E4869.
14. Minaeva E, Forchhammer K & Ermilova E (2015) Glutamine Assimilation and Feedback Regulation of L-acetyl-N-glutamate Kinase Activity in *Chlorella variabilis* NC64A Results in Changes in Arginine Pools. *Protist* **166**(5):493-505.
15. Ermilova E, Lapina T, Zalutskaya Z, Minaeva E, Fokina O & Forchhammer K (2013) PII signal transduction protein in *Chlamydomonas reinhardtii*: localization and expression pattern. *Protist* **164**(1):49-59.

16. Chellamuthu VR, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, Hartmann MD & Forchhammer K (2014) A widespread glutamine-sensing mechanism in the plant kingdom. *Cell* **159**(5):1188-1199.
17. Lapina T, Selim KA, Forchhammer K & Ermilova E (2018) The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins. *Sci Rep* **8**(1):790.
18. Beez S, Fokina O, Herrmann C & Forchhammer K (2009) N-acetyl-L-glutamate kinase (NAGK) from oxygenic phototrophs: P(II) signal transduction across domains of life reveals novel insights in NAGK control. *J Mol Biol* **389**(4):748-758.
19. Heinrich A, Maheswaran M, Ruppert U & Forchhammer K (2004) The *Synechococcus elongatus* P signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. *Mol Microbiol* **52**(5):1303-1314.
20. Huergo LF & Dixon R (2015) The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. *Microbiol Mol Biol Rev* **79**(4):419-35.
21. Schwarz D, Nodop A, Hüge J, Purfürst S, Forchhammer K, Michel KP, Bauwe H, Kopka J & Hagemann M (2011) Metabolic and transcriptomic phenotyping of inorganic carbon acclimation in the Cyanobacterium *Synechococcus elongatus* PCC 7942. *Plant Physiol* **155**(4):1640-55.
22. Llácer JL, Fita I & Rubio V (2008) Arginine and nitrogen storage. *Curr Opin Struct Biol* **18**(6):673-81.
23. Fokina O, Chellamuthu VR, Forchhammer K & Zeth K (2010) Mechanism of 2-oxoglutarate signaling by the Synechococcus elongatus PII signal transduction protein. *Proc Natl Acad Sci U S A* **107**(46):19760-19765.
24. Fokina O, Herrmann C & Forchhammer K (2011) Signal-transduction protein P(II) from *Synechococcus elongatus* PCC 7942 senses low adenylate energy charge in vitro. *Biochem J* **440**(1):147-156.
25. Court SJ, Waclaw B & Allen RJ (2015) Lower glycolysis carries a higher flux than any biochemically possible alternative. *Nat Commun* **6**:8427.
26. Smith DR & Lee RW (2011) Nucleotide Diversity of the Colorless Green Alga *Polytomella parva* (Chlorophyceae, Chlorophyta): High for the Mitochondrial Telomeres, Surprisingly Low Everywhere Els. *J. Eukaryot. Microbiol* **58**(5): 471–473.
27. Vance P & Spalding MH (2005) Growth, photosynthesis, and gene expression in *Chlamydomonas* over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states. *Can J Bot* **83**:796–809.
28. Watzer B, Engelbrecht A, Hauf W, Stahl M, Maldener I & Forchhammer K (2015) Metabolic pathway engineering using the central signal processor PII. *Microb Cell Fact* **14**:192.
29. Maheswaran M, Urbanke C & Forchhammer K (2004) Complex formation and catalytic activation by the PII signaling protein of N-acetyl-L-glutamate kinase from *Synechococcus elongatus* strain PCC 7942. *J Biol Chem* **279**(53):55202-55210.
30. Hauf K, Kayumov A, Gloge F & Forchhammer K (2016) The Molecular Basis of TnrA Control by Glutamine Synthetase in *Bacillus subtilis*. *J Biol Chem* **291**(7):3483-3495.

Supporting Information

Additional supporting information is available in the online version of this article.

Figure Ligands

Fig. 1. Central C- and N- metabolism in nonphotosynthetic alga *P. parva*. (a) Inferred metabolic pathways in non- and photosynthetic algae *P. parva* and *C. reinhardtii*, respectively, with special reference to the TCA- and GS/GOGAT-cycles. Significant metabolic alterations of (b) PEP (C-metabolism), (c) TCA-cycle intermediates, and (d) the major amino acids of N-assimilation reactions and GS/GOGAT-cycle intermediates within the nonphotosynthetic algae *P. parva* cells after shift from rich- (7.5 mM NH_4^+) to low-nitrogen (0.375% mM NH_4^+) conditions in comparison to the photosynthetic algae *C. reinhardtii* under rich-nitrogen (7.5 mM NH_4^+) condition. The metabolites concentrations are relative to *P. parva* cells under high-nitrogen (7.5 mM NH_4^+) supply for three independent replicates, and the standard deviation (SD) is indicated by error bars.

Fig. 2. Multiple sequence alignment for amino acid sequences of PII proteins. Sequences are deduced for PII polypeptides from nonphotosynthetic alga: *Polytomella parva* (Polyt; MMETSP0052_2-20121109|12411), green photosynthetic alga: *Chlamydomonas reinhardtii* (Cr; XP_001703658.1), higher photosynthetic plants: *Physcomitrella patens* (Physco; BAF36548.1), *Arabidopsis thaliana* (At; NP_192099.1), *Oryza sativa Japonica* (Os, Os05g0133100), and *Solanum lycopersicum* (Sl, AAR14689.1), red algae: *Porphyra purpurea* (Pp; NP_053864.1), *Porphyra umbilicalis* (Pu; AFC39923.1), and *Pyropia yezoensis* (Py; AGH27579.1), and bacteria: *Synechococcus elongatus* PCC 7942 (Sy; P0A3F4.1), *Synechocystis* sp. PCC 6803 (Sc; CAA66127.1), and *E. coli* (Ec; CAQ32926.1). The proteins sequences were derived from NCBI database. The regions referring to T-, B- and Q-loops are indicated. Highlighted residues in black are conserved in at least 55% of aligned PIIs proteins. Amino acids in a gray represent similar residues. Boxs I and II indicate PII signature patterns. The positions of known PIIs posttranslational modification sites: the phosphorylation site in cyanobacterial *S. elongatus* PII (S49) and the uridylation site in *E. coli* PII (Y51) are indicated by solid black and white arrows, respectively. The amino acid residues involved in binding of ATP (●), NAGK (■) and 2-OG (▲) are indicated. The lignment were done using the ClustalW program and manually refined.

Fig. 3. Gln-independent *Ppa*PII. (a) The catalytic activity of *Ppa*NAGK with or without *Ppa*PII protein in presence and absence of 5 mM Gln. NAG was used as a variable substrate. The curves were fitted using GraphPad prism program, as indicated. Standard deviations were calculated from independent triplicate experiments. (b) Isolation of *Ppa*PII-*Ppa*NAGK complex using SEC-MALS; SEC-MALS profiles of *Ppa*PII-*Ppa*NAGK complex in presence (black) and

absence (red) of Gln, *Ppa*PII (green), and *Ppa*NAGK (blue), revealed the ability of *Ppa*PII to complex with NAGK independent of Gln. The mass of the eluted particles was determined via MALS and plotted on the right axis. The protein elution profile was monitored using UV signal at 280 nm and plotted on the left. (c) The eluted fraction for *Ppa*PII-*Ppa*NAGK complex shown in (b) was collected and subjected to Glycine-SDS PAGE, indicated the presence of both proteins in the peak, as revealed by a Coomassie-blue stain. See also Figure S3.

Fig. 4. Arginine-feedback inhibition of NAGK activity in the presence and absences of PII. Arginine inhibition of (a) *Ppa*NAGK and (b) *Cr*NAGK in the presence or absence of *Ppa*PII, without or with (5mM) glutamine, as indicated. Arginine inhibition of *Ppa*NAGK in the presence or absence of (c) *Cr*PII and (d) *Os*PII, without or with (5mM) glutamine, as indicated. Data were fitted according to a sigmoidal dose-response curve using a GraphPad Prism, yielding an IC_{50} for arginine. SD as indicated by error bars, represents triplicate independent measurements.

Fig. 5. Effect of glutamine and 2-OG on PII activation of NAGK. The assays were performed in presence of 0.5 and 0.12 mM of arginine for *Ppa*NAGK and *Cr*NAGK, respectively. (a) Glutamine-dependent activation of *Ppa*NAGK and *Cr*NAGK by *Ppa*PII, as indicated. (b) Effect of 2-OG on different PII proteins (*Ppa*PII, *Cr*PII, and *Os*PP) in presence of 5 mM glutamine-mediated activation of *Ppa*NAGK, as indicated. Data were fitted using a GraphPad Prism, yielding an EC_{50} for Gln. SD as indicated by error bars, represents triplicate independent measurements.

Fig. 6. SPR analysis of *Ppa*PII-NAGKs complexes. Independency of (a) *Ppa*PII-*Ppa*NAGK and (B) *Ppa*PII-*Cr*NAGK complexes formation of any tested effectors molecules. Strep-tagged *Ppa*PII with or without effectors was injected on immobilized NAGK (*Ppa*NAGK and *Cr*NAGK) in FC2; the response difference between FC2 and the control FC1 is shown. (a) shows the binding of 1000 nM *Ppa*PII (monomeric concentration) to *Ppa*NAGK in absence of effectors (blue) or in presence of: 2 mM ATP (red), 2 mM ATP + 2 mM Gln (black), 2 mM ATP + 4 mM Gln (dotted-black), 2 mM ADP (green), 2 mM ATP + 2 mM Gln + 4 mM 2-OG (dotted-orange), and 2 mM ATP + 4 mM 2-OG (violet). (b) shows the binding of 1000 nM *Ppa*PII (monomeric concentration) to *Cr*NAGK in absence of effectors (red) or in presence of: 2 mM ATP (dotted-black), 4 mM ADP (blue), 2 mM Gln (gray), 2 mM ATP + 4 mM Gln (orange), 1 mM ATP + 1 mM 2-OG (violet), 2 mM ATP + 10 mM 2-OG (green), 4 mM ADP + 10 mM 2-OG (dotted-red), and 2 mM ATP + 2 mM Gln + 2 mM 2-OG (dotted-blue). (c) The ΔRU for

binding of different concentrations of *Ppa*PII to NAGK without any of effector molecules at t:100s, was used to calculate K_d value for *Ppa*PII. Inset shows the SPR titration of NAGK with increasing concentrations of *Ppa*PII (from 50 nM to 1000 nM monomeric concentrations) in absence of effectors molecules, as indicated.

Fig. 7. SPR analysis of *Cr*PII-*Ppa*NAGKs complex formation. Strep-tagged *Cr*PII with or without effectors was injected on immobilized *Ppa*NAGK in FC2; the response difference between FC2 and the control FC1 is shown. (a) Independency of *Cr*PII-*Ppa*NAGK complex formation of any tested effectors molecules; binding of 1000 nM *Cr*PII (monomeric concentration) to *Ppa*NAGK in absence of effectors (dotted-red) or in presence of: 1 mM ATP (dotted-black), 2 mM ADP (blue), 1 mM ATP + 1 mM Gln (black), 1 mM ATP + 2 mM Gln (red), 1 mM ATP + 3 mM Gln (dotted-blue), 1 mM ATP + 4 mM Gln (green), and 1 mM ATP + 5 mM Gln (violet). (b) binding of 1000 nM *Cr*PII (monomeric concentration) to *Ppa*NAGK in presence of: 2 mM ADP (dotted-black), 2 mM ADP + 2 mM Gln (black), 1 mM ATP (red), 2 mM ATP (blue), and 4 mM ATP (green). (c) No influence of 2-OG on *Cr*PII-*Ppa*NAGK complex formation; binding of 1000 nM *Cr*PII (monomeric concentration) to *Ppa*NAGK in absence of effectors (gray) or in presence of: 2 mM ATP + 3 mM Gln (dotted-black), 2 mM ATP + 3 mM Gln + 1 mM 2-OG (blue), 2 mM ATP + 3 mM Gln + 2 mM 2-OG (red), 2 mM ATP + 3 mM Gln + 4 mM 2-OG (violet), 2 mM ATP + 3 mM Gln + 10 mM 2-OG (black).

Fig. 1

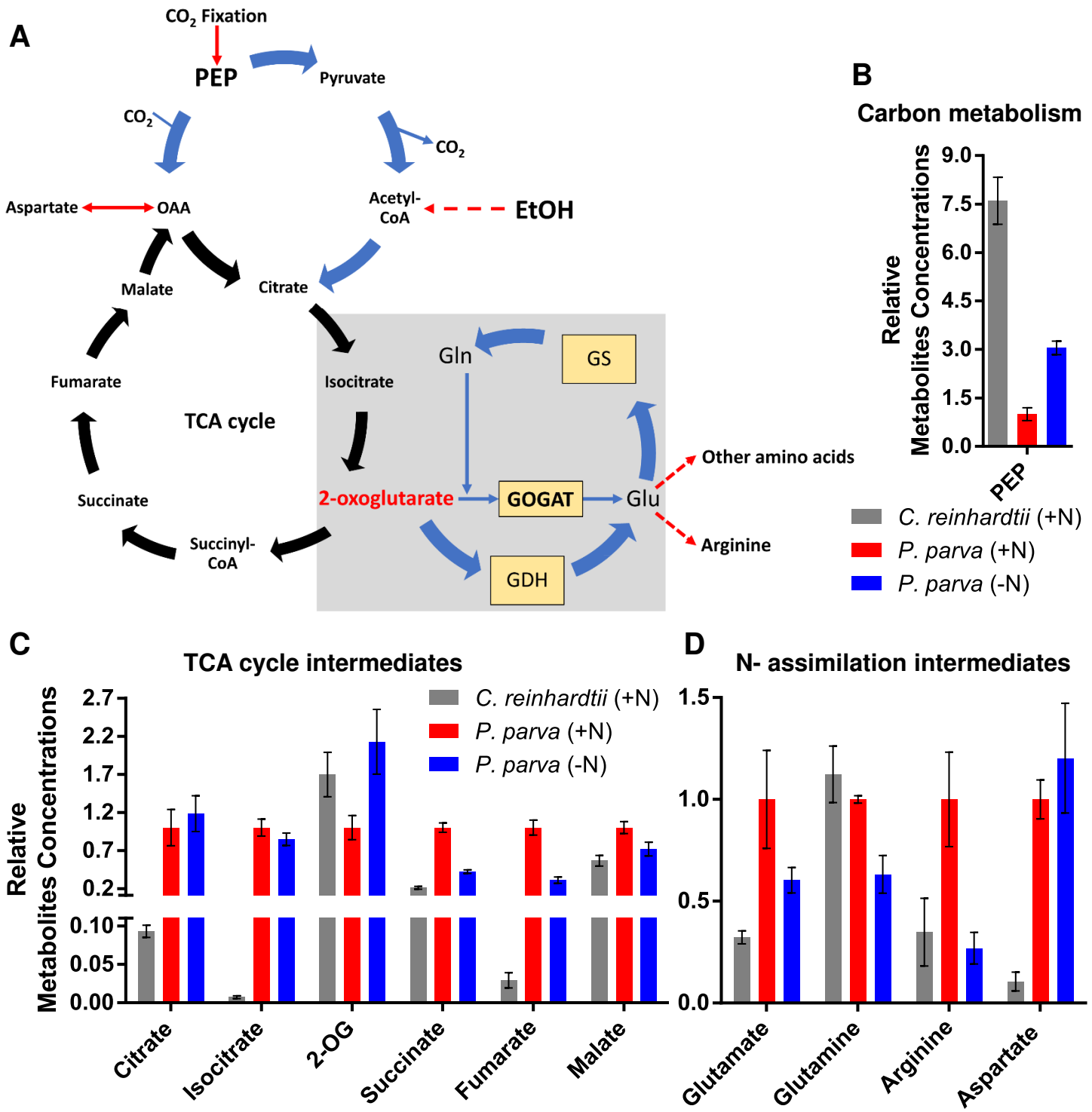

A
CO2 Fixation
PEP
Pyruvate
CO2
CO2
Aspartate
OAA
Acetyl-CoA
EtOH
Malate
Citrate
Fumarate
TCA cycle
Isocitrate
Gln
GS
Succinate
2-oxoglutarate
GOGAT
Glu
Other amino acids
Succinyl-CoA
Arginine
GDH
B
Carbon metabolism
Relative Metabolites Concentrations
9.0
7.5
6.0
4.5
3.0
1.5
0.0
PEP
C. reinhardtii (+N)
P. parva (+N)
P. parva (-N)
C
TCA cycle intermediates
Relative Metabolites Concentrations
2.7
2.2
1.7
1.2
0.7
0.2
0.10
0.05
0.00
Citrate
Isocitrate
2-OG
Succinate
Fumarate
Malate
D
N- assimilation intermediates
1.5
1.0
0.5
0.0
Glutamate
Glutamine
Arginine
Aspartate

Fig. 2

```
Polyt     1 --------------------MNALNFSPTLFNKQTISPTRQTAFITVQGTVRKAKSNFST
Cr        1 -----------------------MALASRTSSAAVVGRSTRSAAVVPVRSIASRCQAARPA
Physco    1 MALQPRLSLSCLRGRSVDACAFAVPASASISASSDACVRIPCWNGASSSSKRLPFFGARV
At        1 ---------------------MAASMTKPISITSLGFYSDRKNIAFSDCISICSGFRHSR
Os        1 ------------------MSSPATAAAAAASCGVLRHHHPPASPRPPPTTTTTTSRLLLA
Sl        1 ---------------------MASPSLSKSNFSLHSFSSSPSLSQFPHFTSITVVQPKFF
Pp        1 ------------------------------------------------------------
Pu        1 ------------------------------------------------------------
Py        1 ------------------------------------------------------------
Sy        1 ------------------------------------------------------------
Sc        1 ------------------------------------------------------------
Ec        1 ------------------------------------------------------------

Polyt    41 SIRQPLVVVSAAKSDAPFKRATYDDLESIKANLSAFPSCEFFRVEAVIRPWRLPFVVEQL
Cr       39 RRASVAVRASDENGSVSVRRAPYAELESIQCDLSAFPGVKFFRIEAIFRPWRLPFVIDTL
Physco   61 ASASDPKSPNWRKRVSGVVQVHLEEYDFDDQSKDYQPSVDFYKVEAVLRPWRLSPVSSAL
At       40 PSCLDLVTKSPSNNSR------VLPVVSAQISSDYIPDSKFYKVEAIVRPWRIQQVSSAL
Os       43 SRSRGLQRPLRVNHAPPRRLPPTAARAQSAAAAGYQPESEFYKVEAILRPWRVPYVSSGL
Sl       40 PSQLTFKR---CQNAP------SFPIIRAQNSPDFVPDAKFYKVEAILRPWRIQQVSSAL
Pp        1 ------------------------------------------MKKIEAIIRPFKLNEVKLAL
Pu        1 ------------------------------------------MKKIEAIIRPFKLNEVKLAL
Py        1 ------------------------------------------MKKIEAIIRPFKLNEVKLAL
Sy        1 ------------------------------------------MKKIEAIIRPFKLDEVKIAL
Sc        1 ------------------------------------------MKKVEAIIRPFKLDEVKIAL
Ec        1 ------------------------------------------MKKIDAIIKPFKLDDVREAL

                                   I
Polyt   101 GNNGIRGMTVTSVHGIGIQGGSRERYGGTEFSQTDLVEKQKVEIVVTRAQANIVSRIIAT
Cr       99 SKYGIRGLTNTPVKGVGVQGGSRERYAGTEFGPSNLVDKEKLDIVVSRAQVDAVVRLVAA
Physco  121 LKMGIRGVTVTDVRGFGAQGGSRERQAGTEYAGDSYLKKVKLEIVVSKDQVEAVIDTIID
At       94 LKIGIRGVTVSDVRGFGAQGGSTERHGGSEFSEDKFVAKVKMEIVVKKDQVESVINTIIE
Os      103 LQMGIRGVTVSDVRGFGAQGGSTERHEGSEFAEDTFIDKVKMEIVVSKDQVEAVVDKIIE
Sl       91 LKMGIRGVTVSDVRGFGAQGGLTERQAGSEFSEDTFVAKVKMEIVVSKDQVEGVIAKIIE
Pp       21 VKGGIGGMTVVKVSGFGRQKGQTERYKGSEYSID-IIDKIKIEIIVSDDKVNSITEIIIK
Pu       21 VKEGIGGMTVVKVSGFGRQKGQTERYKGSEYFID-IIDKIKIEIIISDDKVNKITETIIK
Py       21 VKEGIGGMTVIKVSGFGRQKGQTERYKGSEYSID-IIDKIKIEIIISDDKVEKIVETIIK
Sy       21 VNAGIVGMTVSEVRGFGRQKGQTERYRGSEYTVE-FLQKLKLEIVVEDAQVDTVIDKIVA
Sc       21 VNAGIVGMTVSEVRGFGRQKGQTERYRGSEYTVE-FLQKLKIEIVVDEGQVDMVVDKLVS
Ec       21 AEVGITGMTVTEVKGFGRQKGHTELYRGAEYMVD-FLPKVKIEIVVPDDIVDTCVDTIIR

                II                T-loop      Q-loop
Polyt   161 AAFTGEIGDGKIFIHPVAEVIRIRTAETGFLAEHMAGGMEDMMASKSTA---
Cr      159 SAYTGEIGDGKIFVHPVAEVVRIRTAETGLEAEKMEGGMEDMMKKKK-----
Physco  181 QARTGEIGDGKIFVSPVSDIIRIRTGERGLKAERMAGGRAAMQTSAEGSDGN
At      154 GARTGEIGDGKIFVLPVSDVIRVRTGERGEKAEKMTGDMLSPS---------
Os      163 KARTGEIGDGKIFLIPVSDVIRIRTGERGERAERMAGGLADKLSSAMPIS--
Sl      151 EARTGEIGDGKIFLTPISDVIRVRTGERGEKAERMMGGHADMSSALSTS---
Pp       80 TAKTGEIGDGKIFISDVEQVIRIRTNDLNSAAL-------------------
Pu       80 TAKTGEIGDGKIFISDVEQVIRIRTNDLNSEAL-------------------
Py       80 ASKTGEIGDGKIFISSIERVIRIRTNDLNFEAL-------------------
Sy       80 AARTGEIGDGKIFVSPVDQTIRIRTGEKNADAI-------------------
Sc       80 AARTGEIGDGKIFISPVDSVVRIRTGEKDTEAI-------------------
Ec       80 TAQTGKIGDGKIFVFDVARVIRIRTGEEDDAAI-------------------

                 B-loop                C-loop
```

Fig. 3

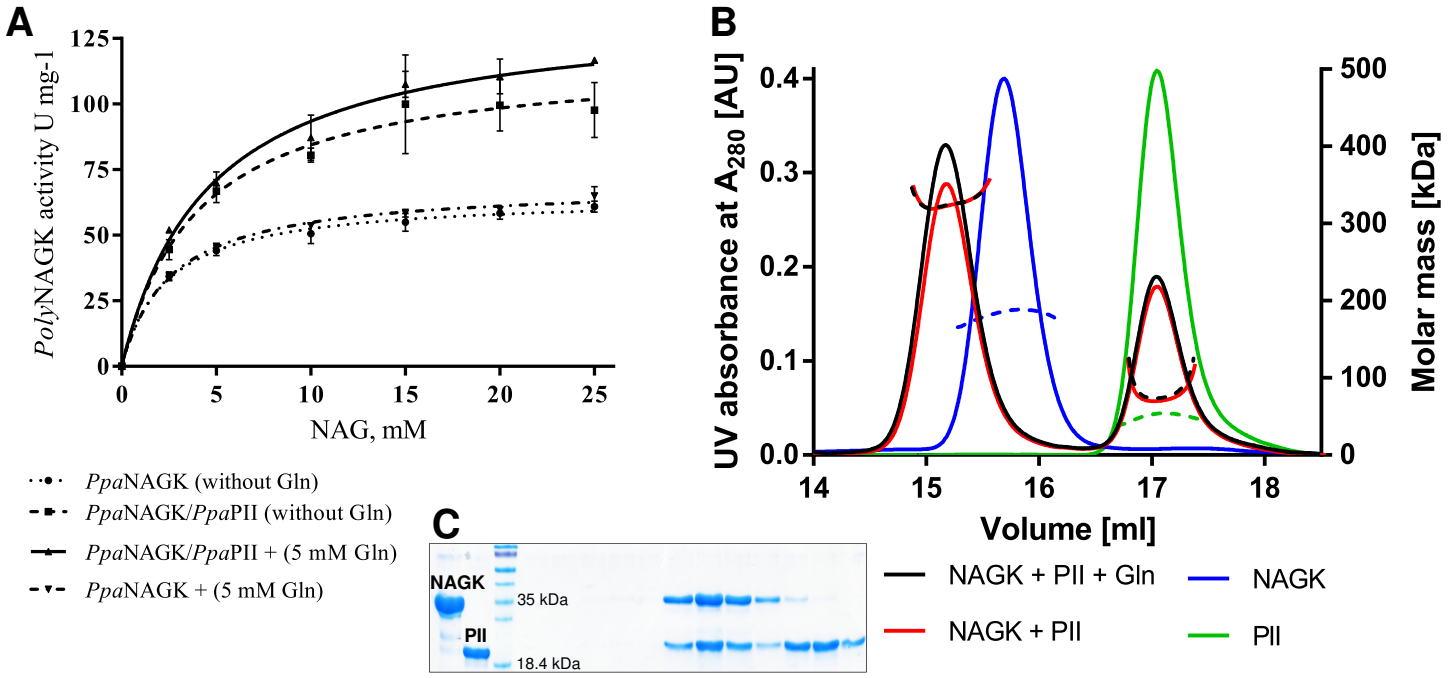

A
PolyNAGK activity U mg-1
125
100
75
50
25
0
0
5
10
15
20
25
NAG, mM
PpaNAGK (without Gln)
PpaNAGK/PpaPII (without Gln)
PpaNAGK/PpaPII + (5 mM Gln)
PpaNAGK + (5 mM Gln)
B
UV absorbance at A280 [AU]
0.4
0.3
0.2
0.1
0.0
Molar mass [kDa]
500
400
300
200
100
0
14
15
16
17
18
Volume [ml]
NAGK + PII + Gln
NAGK
NAGK + PII
PII
C
NAGK
PII
35 kDa
18.4 kDa

Fig. 4

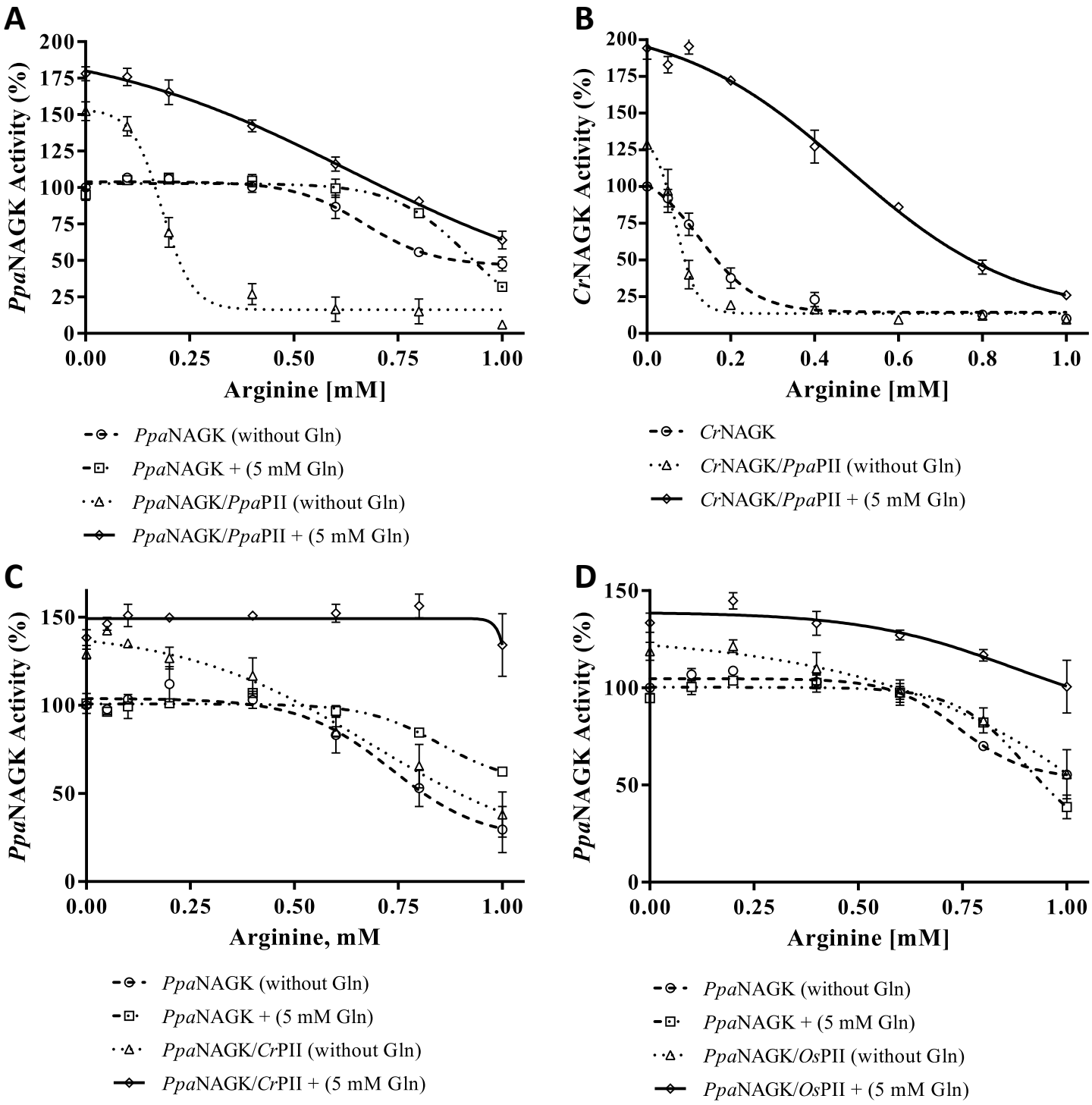

A
PpaNAGK Activity (%)
Arginine [mM]
PpaNAGK (without Gln)
PpaNAGK + (5 mM Gln)
PpaNAGK/PpaPII (without Gln)
PpaNAGK/PpaPII + (5 mM Gln)
B
CrNAGK Activity (%)
Arginine [mM]
CrNAGK
CrNAGK/PpaPII (without Gln)
CrNAGK/PpaPII + (5 mM Gln)
C
PpaNAGK Activity (%)
Arginine, mM
PpaNAGK (without Gln)
PpaNAGK + (5 mM Gln)
PpaNAGK/CrPII (without Gln)
PpaNAGK/CrPII + (5 mM Gln)
D
PpaNAGK Activity (%)
Arginine [mM]
PpaNAGK (without Gln)
PpaNAGK + (5 mM Gln)
PpaNAGK/OsPII (without Gln)
PpaNAGK/OsPII + (5 mM Gln)

Fig. 5

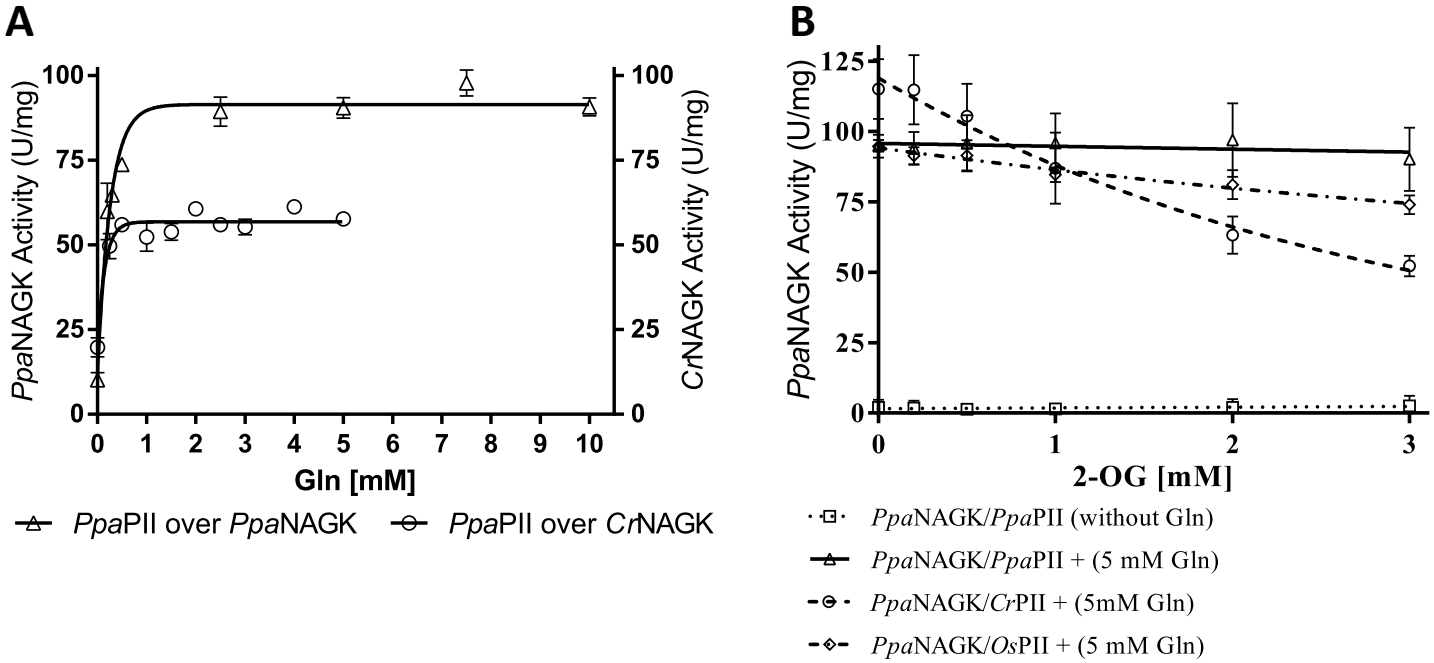

A
PpaNAGK Activity (U/mg)
CrNAGK Activity (U/mg)
Gln [mM]
PpaPII over PpaNAGK
PpaPII over CrNAGK
B
PpaNAGK Activity (U/mg)
2-OG [mM]
PpaNAGK/PpaPII (without Gln)
PpaNAGK/PpaPII + (5 mM Gln)
PpaNAGK/CrPII + (5mM Gln)
PpaNAGK/OsPII + (5 mM Gln)

Fig. 6

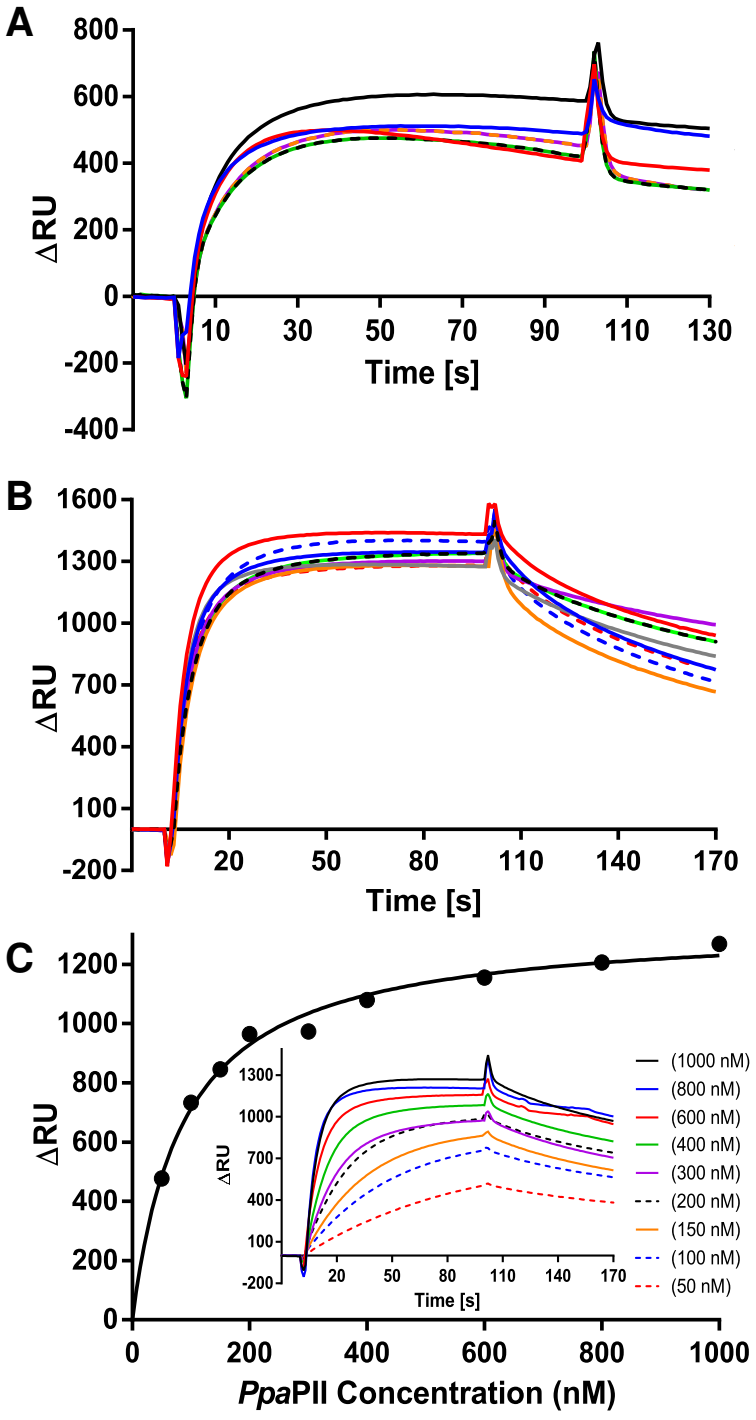
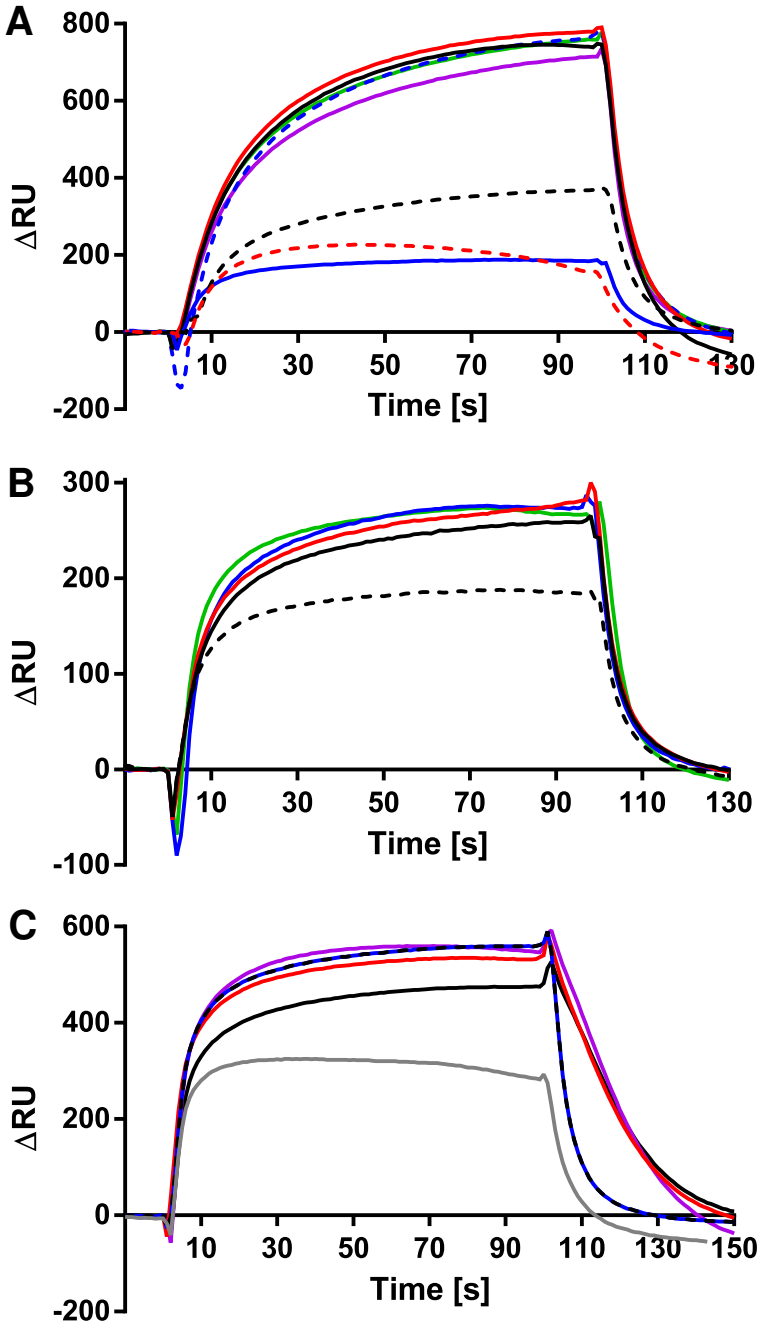

A
800
600
400
200
0
-200
-400
ΔRU
10
30
50
70
90
110
130
Time [s]
B
1600
1300
1000
700
400
100
-200
ΔRU
20
50
80
110
140
170
Time [s]
C
1200
1000
800
600
400
200
0
ΔRU
0
200
400
600
800
1000
PpaPII Concentration (nM)
1300
1000
700
400
100
-200
ΔRU
20
50
80
110
140
170
Time [s]
(1000 nM)
(800 nM)
(600 nM)
(400 nM)
(300 nM)
(200 nM)
(150 nM)
(100 nM)
(50 nM)

Fig. 7


A
800
600
400
200
0
-200
ΔRU
10
30
50
70
90
110
130
Time [s]
B
300
200
100
0
-100
ΔRU
10
30
50
70
90
110
130
Time [s]
C
600
400
200
0
-200
ΔRU
10
30
50
70
90
110
130
150
Time [s]

Supporting Information

Article title: **Interaction of N-acetyl-L-glutamate kinase with the PII signal transducer in the non-photosynthetic alga *Polytomella parva*: Co-evolution towards a hetero-oligomeric enzyme**

Khaled A. Selim[1,*], Tatyana Lapina[2,*], Karl Forchhammer[1] and Elena Ermilova[2]

[1]Organismic Interactions Department, Interfaculty Institute of Microbiology and Infection Medicine Tübingen, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany.

[2]Biological Faculty, Saint-Petersburg State University, Universitetskaya nab. 7/9, Saint-Petersburg 199034, Russia.

[*]Authors contributed equally to this work.

Authors For correspondence:

Karl Forchhammer (karl.forchhammer@uni-tuebingen.de),

Telefon +49 7071 29-72096

Telefax +49 7071 29 - 5843

Elena Ermilova (e.ermilova@spbu.ru)

Tel.: 007 812 4506740

fax: 007 812 4507310.

The following Supporting Information is available for this article:

1. Supplementary Materials and Methods S1
2. Supplementary Table S1
3. Supplementary Figures (S1-S4)
4. Supplementary References

1. Supplementary Materials and Methods S1

1.1. Strains and cultivation conditions.

The whole cloning procedures were performed in *E. coli* NEB 10-beta, while the proteins expression and purification were done using *E. coli* LEMO-21(DE3) and PII-deficient *E. coli* RB9060 (Bueno et al. 1985) in LB medium. The *Polytomella parva* SAG 63-3 culture was obtained kindly from the algal culture collation (SAG-Göttingen University, Germany) as an environmental non-axenic culture. The culture was excessively treated with antibiotics until we were able to isolate a clean axenic culture of *P. parva* SAG 63-3 (Fig. S1). *P. parva* was cultivated in REP media containing 40 mM EtOH as a carbon source and 7.5 mM NH_4Cl as a nitrogen source, pH 4.0 (Atteia et al. 2000) at 22℃ under day/night cycles. The wild-type *Chlamydomonas reinhardtii* CC-125 mt+ [137c] was kindly obtained from Schäffer lab. (ZMBP, Tübingen University), and cultivated in tris-acetate-phosphate (TAP) medium containing 7.5 mM NH_4Cl (Sager and Granick 1954) under day/night cycles at 22℃.

To induce nitrogen deprivation, an exponentially growing culture of *P. parva* under nitrogen rich condition (7.5 mM NH_4Cl) was harvested, washed twice in N-free media, then suspended in fresh media, and divided into two subcultures. One subculture was reinoculated again into nitrogen rich (7.5 mM NH_4Cl) condition, while the other half was reinoculated into nitrogen limiting (0.375 mM NH_4Cl) condition. After 45h, the *P. parva* cultures were harvested to determine the intracellular metabolites using LC-MS, in comparison to standard growing culture of *C. reinhardtii* under nitrogen rich condition (7.5 mM NH_4Cl).

1.2. Metabolite extraction and quantification.

For quantification the intracellular metabolites of 50 ml of exponentially growing cells under the day cycle of *P. parva* under different nitrogen regimes (excess nitrogen of 7.5 mM NH_4Cl or poor nitrogen of 0.375 mM NH_4Cl), and of *C. reinhardtii* under nitrogen rich condition (7.5 mM NH_4Cl) were shock-cooled in ice for 5 min, then rapidly harvested by centrifugation at 4℃. After discarding of liquid media, the cell pellets were immediately frozen in liquid nitrogen. The metabolites extraction and quantification were done according to (Watzer et al. 2015). Briefly, the cells were lyophilized followed by an extraction of the metabolites using a Retsch ball mill (two cycles, 30 sec. each). Extraction was done twice using 400 µl of 80 % methanol containing 0.1 % formic acid followed by a second extraction step with 400 µl of 20 % methanol also containing 0.1 % formic acid. The extracted metabolites combined and concentrated in a speed vac, then dissolved again in 150 µl of 20% methanol containing 0.1 % formic acid (HPLC-grade). Analyses

were done with a Waters UPLC-SynaptG2 LC/MS system. Chromatography was carried out on a 2.1 x 100mm, 1.8µm Waters Acquity HSST3 column. For separation a 10 min gradient from 99 % water to 99% methanol (both solvents with 0.1 % formic acid) was used. The mass spectrometer was operated in ESI negative and positive mode and scanned from 50 to 2000 m/z with a scan rate of 0.5 sec. For the determination of peak areas extracted ion chromatograms were generated and integrated. The quantification of the intracellular metabolites was normalized to cell-dry weight.

1.3. Cloning of *Ppa*PII and *Ppa*NAGK-like proteins.

Gene Blocks, with optimized codon usage for cloning and expression into *E. coli*, encoding for amino acid sequences of mature *Ppa*NAGK-like and *Ppa*PII genes without chloroplast signal peptides, were synthesized by IDT, USA. The first gene Block fragment for the amino acid sequence of the *Ppa*NAGK-like was derived from DNA sequence starting with the 41 amino acid (TSDKK); the gene was amplified using the forward primer 5'-TCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGC-3' and the reverse primer 5'-TATGCTCGAGGATCCGGCTGCTAACAAAGCCCGAAAGGAA-3'. The second gene Block was for the DNA sequence, was derived from the amino acid sequence of the *Ppa*PII starting with the 50 amino acid (SAAKS) and was amplified with the forward primer 5'-AATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATG-3' and the reverse primer 5'-CTGCAGGGGGACCATGGTCTCAGCGCTTGGAGCCACCCGC-3'. Using Gibson assembly, the gene Blocks for *Ppa*NAGK and *Ppa*PII were cloned directly into NdeI-digested pET15b vector (Novagen, Germany) and BsaI-digested pASK-IBA3 vector (IBA, Germany), respectively, as described previously (Gibson et al. 2009). The chloroplast signal peptides were determined using a ChloroP 1.1 Server (http://www.cbs.dtu.dk/services/ChloroP/) (Emanuelsson et al. 1999 and 2007).

1.4. Expression and purification of *Ppa*NAGK, *Cr*NAGK, *Ppa*PII, *Cr*PII and *Os*PII proteins.

The overexpression of the recombinant N-terminal fused His_6-tagged *Ppa*NAGK and *Cr*NAGK were performed in *E. coli* LEMO-21(DE3) and the proteins were affinity purified on a Ni-NTA columns according to (Maheswaran et al. 2004). Overexpression of the recombinant C-terminal fused Strep-tagged PII proteins (*Ppa*PII, *Cr*PII and *Os*PII) were performed in PII-deficient *E. coli* RB9060 (Bueno et al. 1985) and the proteins were affinity purified on a Strep-Tactin II column according to Heinrich et al. (2004) and Selim et al. (2018).

1.5. Coupled NAGK activity assay.

The activity of NAGK was assessed using a coupled enzyme assay in which the production of ADP after the consumption of ATP for phosphorylation of NAG was associated with the oxidation of NADH by pyruvate kinase and lactate dehydrogenase as described previously (Beez et al. 2009, Lapina et al. 2018). The standard reaction mixture composed of 50 mM imidazole pH 7.5, 50 mM KCl, 20 mM $MgCl_2$, 0.4 mM NADH, 1 mM phosphoenolpyruvate, 5 mM ATP, 0.5 mM DTT, 11 U lactate dehydrogenase, 15 U pyruvate kinase and 50 mM NAG and the reaction was started by the addition of 1.5 µg NAGK. When necessary, PII protein was added to the reaction mix in equimolar concentration. When needed, the effector molecules 2-OG, Gln and Arg were added to the reaction mixtures at concentrations as indicated. The oxidation of NADH was measured at 340 nm for 10 min with SPECORD-spectrophotometer (model-210 PLUS, Analytik Jena AG). One molecule oxidation of NADH is proportional to one molecule phosphorylation of NAG. One unit of NAGK catalysis the conversion of 1 µmol of NAG min^{-1}, calculated with the molar absorption coefficient of NADH of 6178 L mol^{-1} cm^{-1} for at 340 nm. Means of triplicate experimental determinations are shown with a standard deviation of less than 5%. The enzymatic constants K_m, k_{cat}, IC_{50} and EC_{50} were calculated from the velocity slopes using the Graph-Pad Prism software program (GraphPad Software, USA).

1.6. Surface plasmon resonance spectroscopy analysis (SPR spectroscopy).

SPR experiments were done at 25℃ using a BIAcore-X biosensor system (Biacore AB, Uppsala, Sweden) in HBS buffer (10 mM HEPES, 150 mM NaCl, 2 mM $MgCl_2$ and 0.005% NP-40, pH 7.5) with a flowrate of 15µl/min, as described previously (Maheswaran et al. 2004, Chellamuthu et al. 2014). The recombinant His_6-tagged NAGKs (*Ppa*NAGK and *Cr*NAGK) proteins were immobilized on the flow cell (FC2) of the Ni^{2+}-loaded NTA-biosensor chip. NAGKs in HBS buffer were injected (50µl) until a saturation of NTA-biosensor chip by a signal of ≈ 3000-4000 resonance units (RU), which corresponds to a surface concentration change of 3-4 (ng/mm^2). To evaluate the effect of the effector molecules on the PII-NAGK complex formation for the binding of PII (*Ppa*PII, *Cr*PII, and *Os*PII) proteins to the immobilized His6-tagged NAGKs, the strep-tagged PII proteins (100-1000 nM) as indicated in HBS buffer, were incubated in ice for 5min with/without different combinations of the effector molecules (as indicated). Followed by injection of PII proteins (25 µl) into both FC1 (control, for unspecific binding of PIIs to the surface of the sensor chip) and FC2 (NAGKs immobilized) on the sensor chip. The specific binding of PIIs to NAGKs was recorded as the difference in the response signal of FC2-FC1 (ΔRU). *Cr*PII protein dissociates immediately after the end of the injection, making immobilized NAGKs ready for the next injection. Whereas,

*Ppa*PII protein form a strong complex with NAGKs and dissociates very slowly over the time. In case of injection of *Ppa*PII protein, to refresh the NTA sensor chip for another run, 25 µl of 1 M imidazole pH 7.0 was injected to remove the immobilized NAGKs-bound to *Ppa*PII. To regenerate the NTA sensor chip, 50 µl of 0.4 M EDTA pH 7.5 was injected to remove the loaded Ni^{2+} and His-tagged NAGKs. Subsequently, the sensor chip could be reloaded once again with Ni^{2+} and fresh NAGKs as described. The regeneration procedure was done when the response of PII binding to the immobilized NAGK started to decrease.

1.7. Size Exclusion Chromatography and Multiangle Light Scattering (SEC-MALS) Analysis.

Analytical size exclusion chromatography was carried out according to (Hauf et al. 2016) on a micro-Äkta purifier system equipped with Superose 6 Increase 10/300 GL (GE Healthcare, column volume of 24 ml). The Superose column was coupled to a triple-angle light scattering (MALS) detector: miniDAWN™ TREOS® system (Wyatt Technology Corp., CA) and a refractometer (Optilab T-rEX, Wyatt). The column calibration was done using standard proteins: thyroglobulin (670 kDa), ferritin (440 kDa), globulin (158 kDa), conalbumin (75 kDa), ovalbumin (44 kDa), carbonic anhydrase (29 kDa), RNase (13.7 kDa) (Bio-Rad gel filtration standard, GE Healthcare LMW gel filtration calibration kit). The protein, Bovine serum albumin (BSA) was used to calibrate and validate the MALS analysis. The running buffer consisted of 10 mM Tris pH 7.8, 300 mM NaCl, 2 mM $MgCl_2$, 0.02% NaN_3 and 2% glycerol. The samples were centrifuged for 5 min at 14,000 rpm, and 100µl of the supernatant were injected for analysis with a flow rate 0.5 ml/min. The resulting data were analyzed with ASTRA program (Wyatt Technology). The experiments were performed at room temperature, and the apparent molecular weight was calculated from an average of two injections. The elution volume was plotted against the UV signal and molecular weight profiles. The apparent molecular weights were derived from MALS data. The chromatographic elution profiles were collected (0.5 ml fractions) and analyzed by Glycine-SDS PAGE.

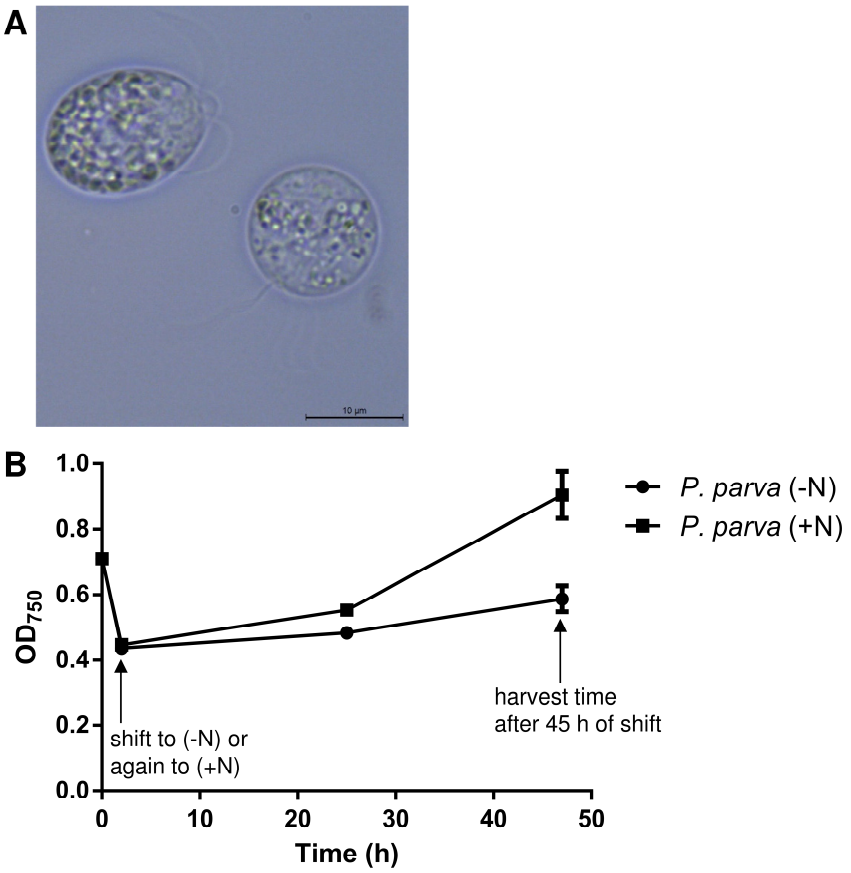
2. Supplementary Table S1

Table S1. List of identified metabolites by LC-MS normalized to 1 mg of algal cell dry weight including standard deviation (SD) of three biological replicates for *Polytomella parva* (under nitrogen excess and limiting conditions) and *Chlamydomonas reinhardtii* (under nitrogen rich conditions).

Compound/ Organism (condition)	*C. reinhardtii* (Nitrogen rich condition, 7.5 mM NH_4^+)		*P. parva* (Nitrogen rich condition, 7.5 mM NH_4^+)		*P. parva* (Nitrogen limiting condition, 0.375 mM NH_4^+)	
	Area of MS base-peak normalized to 1 mg cell dry weight	**SD**	**Area of MS base-peak normalized to 1 mg cell dry weight**	**SD**	**Area of MS base-peak normalized to 1 mg cell dry weight**	**SD**
Phosphoenol pyruvate (PEP)	6.04	0.58	0.79	0.16	2.42	0.17
Citrate	72.62	6.11	780.45	187.29	924.38	183.54
α-ketoglutarate (2-OG)	22.04	3.79	12.99	2.08	27.64	5.53
Succinate	111.06	9.93	531.39	33.61	224.57	12.76
Fumarate	1.75	0.59	59.56705	5.93	18.48	2.40
Malate	269.77	33.34	476.80989	37.18	342.47	43.39
Isocitrate	0.83	0.21	113.8837	12.57	96.47	9.22
Glutamate	45.58	4.51	141.31261	33.92	85.19	8.87
Aspartate	7.88	3.42	74.238451	7.11	89.25	19.99
Glutamine	2.72	0.34	2.4224576	0.05	1.53	0.22
Arginine	20.34	9.74	58.420397	13.58	15.71	4.54

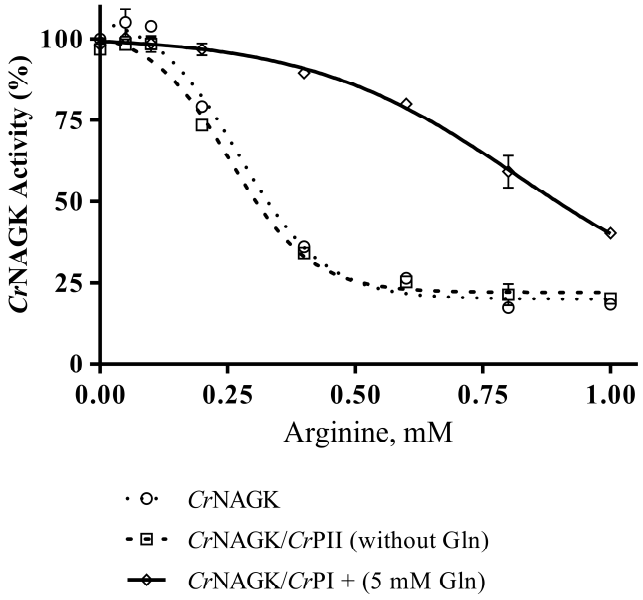
3. Supplementary Figures

Supplementary Fig. S1

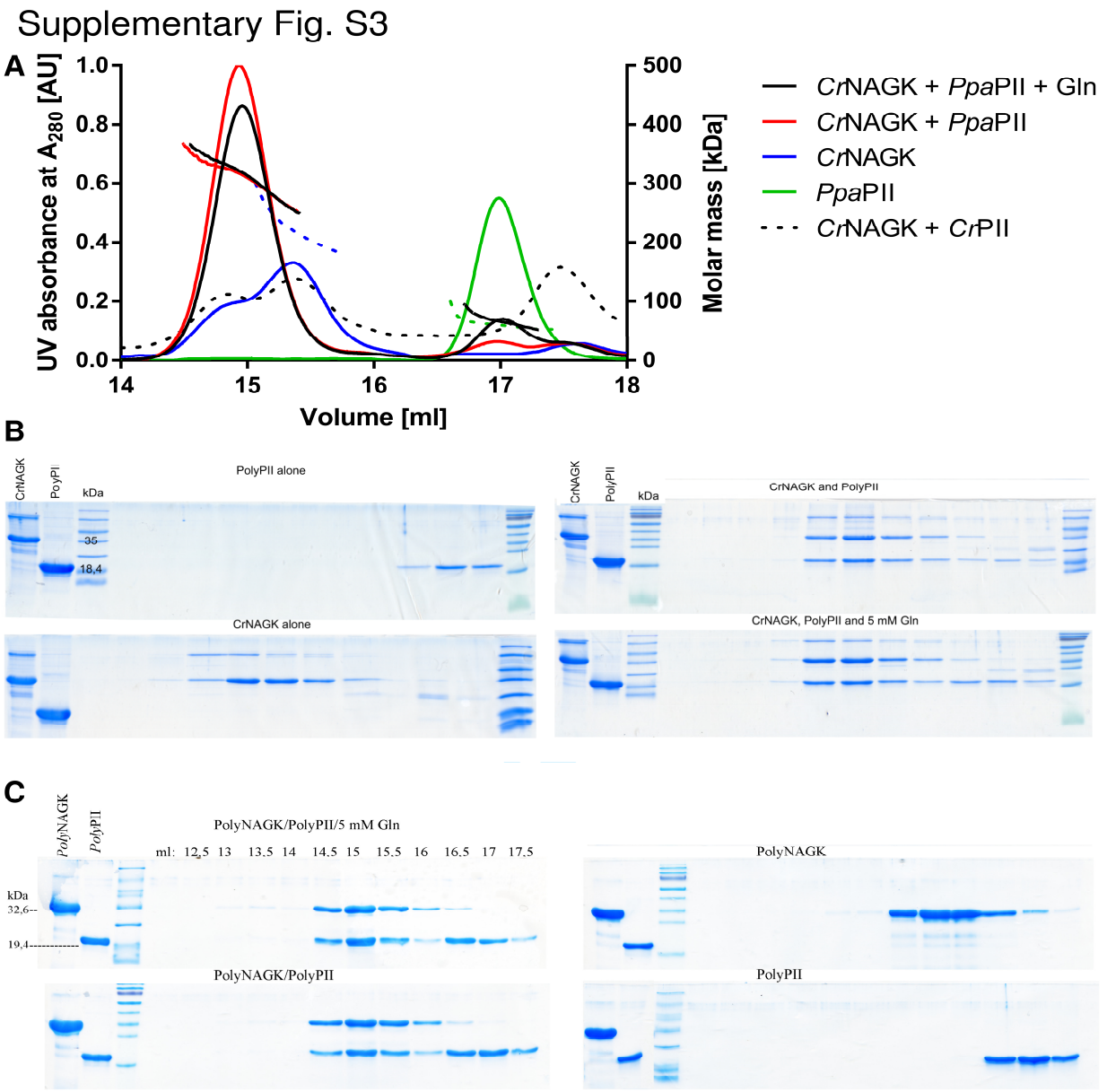


Supplementary Fig. S1. Growth of axenic culture of *Polytomella parva* SAG 63-3. (A) Microscopic examination of isolated axenic culture of *P. parva* SAG 63-3. Bar, 10 µm (B) Growth of *P. parva* under N-limited and rich conditions, showing the time point (45 h) for harvesting *P. parva* for metabolomic analysis after the shift. In the beginning of the experiment, exponentially growing culture of *P. parva* under nitrogen-rich conditions was collected and shifted to N-limiting (0.375 mM NH_4^+) or back again to the N-rich conditions (7.5 mM NH_4^+).

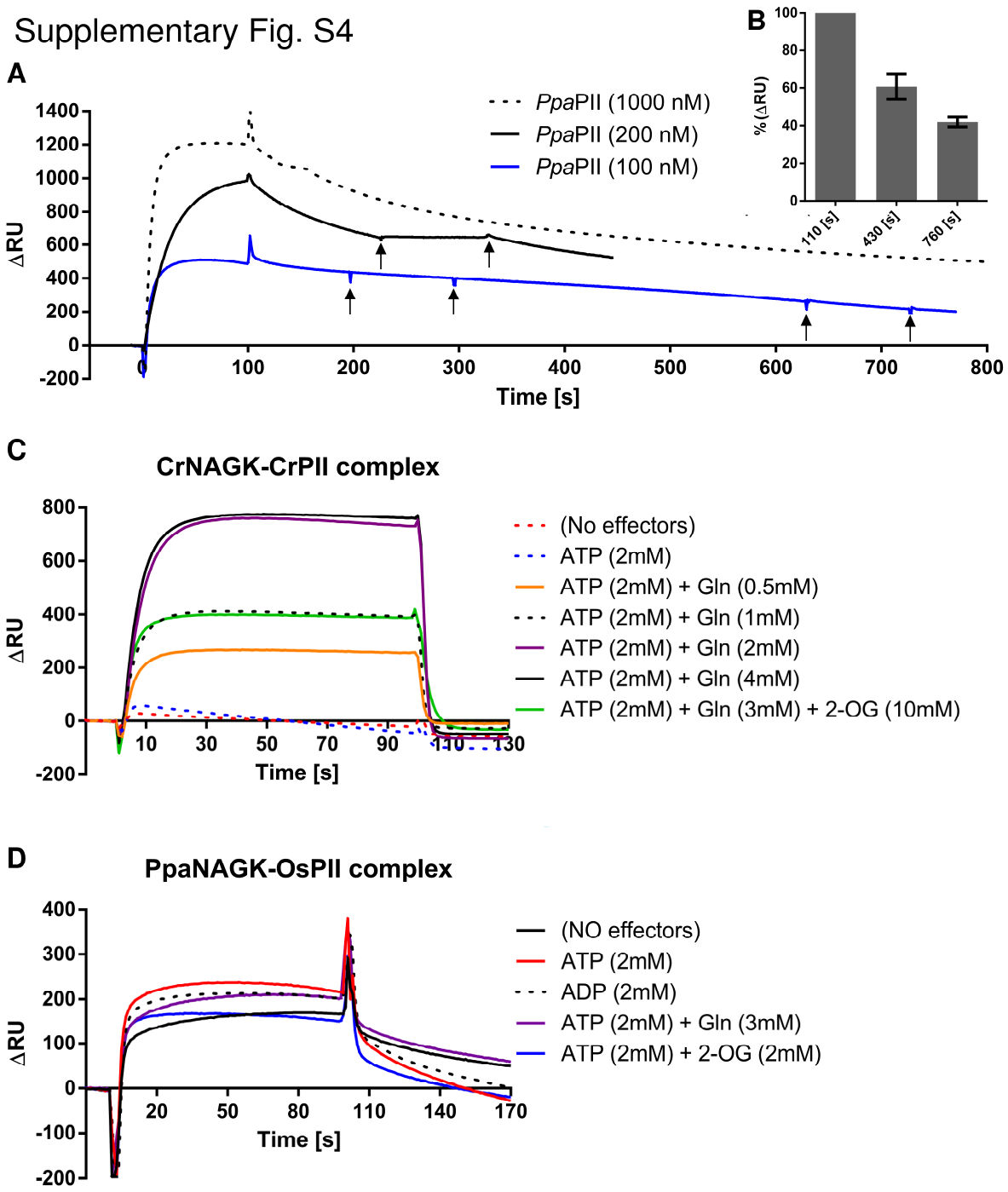
Supplementary Fig. S2



Supplementary Fig. S2. Arginine feedback inhibition of *Cr*NAGK activity in the presence of *Cr*PII. The activity was measured in the presence or absence of 5 mM Gln in the presence or absence of CrPII. Data were fitted to a sigmoidal dose-response curve using GraphPad Prism, yielding an EC_{50} (arginine) of 0.27 ± 0.02 mM, 0.25 ± 0.01 and 0.82 ± 0.09 mM for free *Cr*NAGK, *Cr*PII-*Cr*NAGK in absence of Gln, and *Cr*PII-*Cr*NAGK in presence of Gln, respectively. The obtained data were comparable with the previously published results by (Chellamuthu et al. 2014).



Supplementary Fig. S3. Isolation of *Ppa*PII and NAGKs (*Ppa*NAGK and *Cr*NAGK) complexes using SEC-MALS. (A) SEC-MALS profiles of *Ppa*PII and *Cr*NAGK complex in presence (black) and absence (red) of Gln, PpaPII (green), Uncomplexed *Cr*PII and *Cr*NAGK (dotted-black) and *Cr*NAGK (blue), revealed the ability of *Ppa*PII to complex with *Cr*NAGK independent of Gln. The mass of the eluted particles was determined via MALS and plotted on the right axis. The protein elution profile was monitored using UV signal at 280 nm and plotted on the left. (B & C) The eluted fractions shown in supplementary (fig. S2A) and fig. (3B) were collected and subjected to Glycine-SDS PAGE; indicated the presence of both proteins in the complex peaks of *Ppa*PII-*Cr*NAGK and *Ppa*PII-*Ppa*NAGK, as revealed by a Coomassie-blue stain.



Supplementary Fig. S4. SPR analysis for PII-NAGK complex formation. (A) *Ppa*PII-*Ppa*NAGK complex formation revealed insensitivity of the complex with further injection of 2 mM ADP (indicated by arrows) and strong stability with very slow dissociation rate of the complex over 700 sec after the finish of *Ppa*PII injection, as indicated. (B) Stability of *Ppa*PII-*Ppa*NAGK complex; shows the average of the response signal in % at *t:430s* (330s after the end of the injection) and at *t:760s* (660s after the end of the injection) of *Ppa*PII-*Ppa*NAGK complex in

absence of any effector molecules. The signal at *t:110s* (10s after the end of the injection) was normalized to 100%. The very slow dissociation over the time is an indicator for the stability of the complex. (C) Strict Mg-ATP/Gln dependency of *Cr*PII binding to *Cr*NAGK, as indicated. (D) The ability *Ppa*NAGK to complex with *Os*PII independent of any effector molecules, with no negative influence of ADP and 2-OG, as indicated.

4. Supplementary References

1. Atteia A, van Lis R, Ramírez J, González-Halphen D. **2000**. *Polytomella* spp. growth on ethanol. Extracellular pH affects the accumulation of mitochondrial cytochrome c550. *Eur J Biochem*. 267(10):2850-2858.
2. Beez S, Fokina O, Herrmann C, Forchhammer K. **2009**. N-acetyl-L-glutamate kinase (NAGK) from oxygenic phototrophs: P(II) signal transduction across domains of life reveals novel insights in NAGK control. *J Mol Biol*. 389(4):748-758.
3. Bueno R, Pahel G, Magasanik B. **1985**. Role of glnB and glnD gene products in regulation of the glnALG operon of *Escherichia coli*. *J Bacteriol*. 164(2):816-822.
4. Chellamuthu VR, Ermilova E, Lapina T, Lüddecke J, Minaeva E, Herrmann C, Hartmann MD, Forchhammer K. **2014**. A widespread glutamine-sensing mechanism in the plant kingdom. *Cell*. 159(5):1188-1199.
5. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. **2007**. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2(4):953-71.
6. Emanuelsson O, Nielsen H, von Heijne G. **1999**. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci*. 8(5):978-84.
7. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA 3rd, Smith HO. **2009** Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*. 6(5):343-345.
8. Hauf K, Kayumov A, Gloge F, Forchhammer K. **2016**. The Molecular Basis of TnrA Control by Glutamine Synthetase in Bacillus subtilis. *J Biol Chem*. 291(7):3483-3495.
9. Heinrich A, Maheswaran M, Ruppert U, Forchhammer K. **2004**. The *Synechococcus elongatus* P signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. *Mol Microbiol*. 52(5):1303-1314.
10. Lapina T, Selim KA, Forchhammer K, Ermilova E. **2018**. The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins. *Sci Rep*. 8(1):790.
11. Maheswaran M, Urbanke C, Forchhammer K. **2004**. Complex formation and catalytic activation by the PII signaling protein of N-acetyl-L-glutamate kinase from *Synechococcus elongatus* strain PCC 7942. *J Biol Chem*. 279(53):55202-55210.
12. Sager R, Granick S. **1954**. Nutritional control of sexuality in *Chlamydomonas reinhardii*. *J Gen Physiol*. 37(6):729-742.
13. Selim KA, Haase F, Hartmann MD, Hagemann M, Forchhammer K. **2018**. P_{II}-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. *Proc Natl Acad Sci U S A*. 115(21):E4861-E4869.
14. Watzer B, Engelbrecht A, Hauf W, Stahl M, Maldener I, Forchhammer K. **2015**. Metabolic pathway engineering using the central signal processor PII. *Microb Cell Fact*. 14:192.

<u>Publication 4:</u> in BBA Bioenergetics; 1860(6):519-532.

doi: 10.1016/j.bbabio.2019.04.007.

Walter J, **<u>Selim KA</u>**, Leganés F, Fernández-Piñas F, Vothknecht UC, Forchhammer K, Aro E-M, Gollan PJ. (2019) A novel Ca^{2+}-binding protein influences photosynthetic electron transport in *Anabaena* sp. PCC 7120. Biochim Biophys Acta Bioenerg. 1860(6):519-532.

Elsevier Editorial System(tm) for BBA - Bioenergetics

Manuscript Draft

Manuscript Number: BBABIO-19-4

Title: A novel Ca2+-binding protein influences photosynthetic electron transport in Anabaena sp. PCC 7120

Article Type: Regular Paper

Keywords: calcium; EF-hand; Anabaena; cyanobacteria; photosynthesis; phycobilisomes

Corresponding Author: Dr. Peter Gollan,

Corresponding Author's Institution:

First Author: Julia Walter

Order of Authors: Julia Walter; Khaled A Selim; Francisco Leganés; Francisca Fernández-Piñas; Ute C Vothknecht; Karl Forchhammer; Eva-Mari Aro; Peter Gollan

Abstract: Ca2+ is a potent signalling molecule that regulates many cellular processes. In cyanobacteria, Ca2+ has been linked to cell growth, stress response and photosynthesis, and to the development of specialist heterocyst cells in certain nitrogen-fixing species. Despite this, the pathways of calcium signal transduction in cyanobacteria are poorly understood, and only a few protein components are known. The current study describes a previously unreported calcium-binding protein which was called the Calcium Sensor EF-hand (CSE), which is conserved in filamentous, nitrogen-fixing cyanobacteria. CSE is shown to bind calcium, which induces a conformational change in the protein structure. Poor growth of a strain of Anabaena sp. PCC 7120 overexpressing CSE was attributed to diminished photosynthetic performance. Transcriptomics, biophysics and proteomics analyses revealed modifications in the light-harvesting phycobilisome and photosynthetic reaction center- protein complexes, and downregulated respiration.

Suggested Reviewers: Anthony Larkum
University of Technology, Sydney
Anthony.Larkum@uts.edu.au
Expertise in photosynthesis and photoinhibition in cyanobacteria, and in ion biochemistry

Masahiko Ikeuchi
University of Tokyo
mikeuchi@bio.c.u-tokyo.ac.jp
Expertise in cyanobacterial thylakoid protein complexes, light-harvesting

Jingdong Zhao
Peking University, Beijing
jzhao@pku.edu.cn
Expertise in Ca2+ signalling and light-harvesting in cyanobacteria

Jens Appel

Helmholtz Center for Environmental Research
jens.appel@ufz.de
Expertise in photosynthetic electron transport and cyanobacterial metabolism

Yumiko Sakuragi
University of Copenhagen
ysa@plen.ku.dk
Expertise in cyanobacterial biotechnology and metabolism

14 January 2019

Dear Editor,

I am pleased to submit a manuscript reporting our original research, entitled “A novel Ca2+-binding protein influences photosynthetic electron transport in Anabaena sp. PCC 7120” to BBA Bioenergetics. Herein, we describe a previously unreported EF-hand protein occurring in filamentous, nitrogen-fixing cyanobacteria. The protein demonstrates selective Ca2+-binding activity, and has a deleterious impact on the structure and function of protein complexes of the photosynthetic membrane when over-expressed in Anabaena sp. PCCC 7120.

We believe this manuscript would be a good fit for publication in BBA Bioenergetics, as our study touches on several key areas of the journal, including photosynthetic membrane bioenergetics, protein structure and cell signalling. Furthermore, we believe the novel EF-hand protein described here will attract readers from a range of interests, including photosynthesis and Ca2+ signalling in both prokaryotic and eukaryotic species.

Thank you for considering our manuscript for publication in BBA Bioenergetics.

With best regards,

Peter Gollan

peter.gollan@utu.fi

Molecular Plant Biology
Department of Biochemistry
University of Turku, Finland

A novel Ca^{2+}-binding protein influences photosynthetic electron transport in *Anabaena* sp. PCC 7120

Julia Walter[1], Khaled A. Selim[2,3], Francisco Leganés[4], Francisca Fernández-Piñas[4], Ute C. Vothknecht[5], Karl Forchhammer[3], Eva-Mari Aro[1], Peter J. Gollan[1]

[1] Department of Biochemistry, Molecular Plant Biology, University of Turku, Tykistökatu 6A, 20520 Turku, Finland

[2] Protein Evolution Department, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, 72076 Tübingen, Germany

[3] Interfaculty Institute of Microbiology and Infection Medicine/Organismic Interactions Department, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

[4] Departamento de Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, Calle Darwin 2, 28049 Madrid, Spain

[5] Institute for Cellular and Molecular Botany, Plant Cell Biology, University of Bonn, Kirschallee 1, 53115 Bonn, Germany

J.W. julwal@utu.fi

K.A.S. khaled.selim@uni-tuebingen.de

F.L. Francisco.leganes@uam.es

F.F.P. francisca.pina@uam.es

U.C.V. vothknecht@uni-bonn.de

K.F. karl.forchhammer@uni-tuebingen.de

E.M.A. evaaro@utu.fi

P.J.G. peter.gollan@utu.fi (corresponding author)

A novel Ca^{2+}-binding protein influences photosynthetic electron transport in *Anabaena* sp. PCC 7120

Julia Walter[1], Khaled A. Selim[2,3], Francisco Leganés[4], Francisca Fernández-Piñas[4], Ute C. Vothknecht[5], Karl Forchhammer[3], Eva-Mari Aro[1], Peter J. Gollan[1]

[1] Department of Biochemistry, Molecular Plant Biology, University of Turku, Tykistökatu 6A, 20520 Turku, Finland

[2] Protein Evolution Department, Max Planck Institute for Developmental Biology, Max-Planck-Ring 5, 72076 Tübingen, Germany

[3] Interfaculty Institute of Microbiology and Infection Medicine/Organismic Interactions Department, Eberhard-Karls-Universität Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

[4] Departamento de Biología, Facultad de Ciencias, Universidad Autónoma de Madrid, Calle Darwin 2, 28049 Madrid, Spain

[5] Institute for Cellular and Molecular Botany, Plant Cell Biology, University of Bonn, Kirschallee 1, 53115 Bonn, Germany

J.W. julwal@utu.fi

K.A.S. khaled.selim@uni-tuebingen.de

F.L. Francisco.leganes@uam.es

F.F.P. francisca.pina@uam.es

U.C.V. vothknecht@uni-bonn.de

K.F. karl.forchhammer@uni-tuebingen.de

E.M.A. evaaro@utu.fi

P.J.G. peter.gollan@utu.fi (corresponding author)

Summary

Ca^{2+} is a potent signalling molecule that regulates many cellular processes. In cyanobacteria, Ca^{2+} has been linked to cell growth, stress response and photosynthesis, and to the development of specialist heterocyst cells in certain nitrogen-fixing species. Despite this, the pathways of calcium signal transduction in cyanobacteria are poorly understood, and only a few protein components are known. The current study describes a previously unreported calcium-binding protein which was called the Calcium Sensor EF-hand (CSE), which is conserved in filamentous, nitrogen-fixing cyanobacteria. CSE is shown to bind calcium, which induces a conformational change in the protein structure. Poor growth of a strain of *Anabaena* sp. PCC 7120 overexpressing CSE was attributed to diminished photosynthetic performance. Transcriptomics, biophysics and proteomics analyses revealed modifications in the light-harvesting phycobilisome and photosynthetic reaction center- protein complexes, and downregulated respiration.

Keywords: calcium; EF-hand; Anabaena; cyanobacteria; photosynthesis; phycobilisomes

1. Introduction

Calcium ions (Ca^{2+}) play a pivotal role in a variety of cellular processes through its capacity to bind to proteins, changing their shape and charge [1]. A high intracellular Ca^{2+} concentration ($[Ca^{2+}]_i$) can be toxic because free Ca^{2+} can precipitate phosphate ions, which are essential for metabolic processes. Therefore $[Ca^{2+}]_i$ is closely monitored and tightly regulated by Ca^{2+} channels and pumps, and Ca^{2+}-binding proteins [2,3]. In plants, free Ca^{2+} in the cytoplasm is maintained at around 100 nM, while higher concentrations are sequestered in various organelles. Rapid changes in the concentration of free Ca^{2+} encode signals that regulate numerous stress and developmental processes [reviewed in 4,5]. In the chloroplast, hormone signalling, photosynthesis and CO_2 fixation are regulated by changes in Ca^{2+} concentration [reviewed in 6,7]. In cyanobacteria, which represent the photosynthetic ancestors of plant chloroplasts [8], Ca^{2+} was found to stimulate intracellular pH homeostasis in low external pH, thus preventing acidification of the cytoplasm and protecting physiological processes such as growth, photosynthesis and nitrogen (N) fixation from inhibition [9]. A Ca^{2+} signal observed in cyanobacteria during light-to-dark transitions [10], similar to the Ca^{2+} transient in plant chloroplasts that is induced by darkness [11], was partially inhibited by a Ca^{2+} channel blocker or calmodulin inhibitor. Hence, Ca^{2+} may be taken up in cyanobacteria from the extracellular medium by Ca^{2+}-binding proteins or a Ca^{2+} pump, upon changes in the redox state of the plastoquinone (PQ) pool [10].

The major thylakoid protein complexes responsible for transformation of sunlight energy into chemical energy (ATP and NADPH) are well conserved between plants and cyanobacteria. Photosystem II (PSII), photosystem I (PSI) and the cytochrome b_6f complex (cyt b_6f) operate in series to transport electrons and to create a proton motive force that drives the ATP synthase complex. Ca^{2+} is essential for the water-splitting activity of the PSII oxygen-evolving complex [12-16], which liberates electrons from water to be used for CO_2 fixation and concomitantly releases O_2 to the atmosphere. PSII crystal structures of *Thermosynechococcus elongatus* and *T. vulgaris* showed that PSII monomers contain four Ca^{2+} ions, one of which is part of the oxygen-evolving complex and others are ligands in CP43 and PsbK [17,18]. PSI also binds calcium. Six Ca^{2+} ions were detected in the crystal structure of trimeric PSI in the *Synechocystis* sp. PCC 6803 PSI complex, two of which were localised to PsaL and PsaB subunits, suggesting a role in the oligomerisation of PSI or in providing a binding site for an unidentified regulatory protein [19].

While photosynthetic processes remain highly conserved between cyanobacteria and plants, one major difference concerns the capturing of photons. In plants, light energy is collected by the membrane-embedded light-harvesting complexes, which channel excitation energy towards the reaction centres of the photosystems. In cyanobacteria, however, light-harvesting antennae complexes, called phycobilisomes (PBS), are bound to the photosystems at the stromal face of the

thylakoid membrane, with a special type of PBS being connected to PSI [20]. In *Synechocystis* sp. PCC 6803, even a megacomplex composed of PBS-PSII-PSI has been reported [21]. PBS are composed of different phycobilin pigment-binding proteins (PBP) and linker proteins, which are organised in allophycocyanin (APC) core cylinders connected to the photosystem reaction centres, and peripheral rod-shaped antennae linked to the APC core. The rods consist of core-connected phycocyanin (PC) discs, while some strains have additional phycoerythrocyanin (PEC) discs at the distal end of each rod [reviewed in 22]. Depending on the cyanobacterial species, the PBS composition and structure can differ in the number and type of PBPs, the type of bound chromophores (pigments) and the number of rods and core cylinders per PBS. In the multicellular model organism *Nostoc* sp. PCC 7120 (herein referred to as *Anabaena*), the APC core contains five cylinders from which eight rods radiate. APC, PC and PEC bind different numbers of the chromophore phycocyanobilin (PCB), resulting in different spectral features for each PBP. PEC absorbs light of shorter wavelengths (absorption maximum at 570 nm), whereas APC absorbs light of longer wavelengths (absorption maximum at 650 nm) [23]. Excitation energy is transferred from PEC via PC and APC to the terminal emitter ApcE, which is a pigmented core-membrane linker protein (L_{CM}) connected to the PSII reaction centre in close proximity to CP43 [21,24,25]. Similarly, the terminal emitter ApcD connects PBS to PSI via hydrophobic interactions at the interface of two PSI monomers [20,21,26].Other linker proteins that do not bind pigments connect the discs, rods and cores within the PBS, to form complexes of around 6 000 kDa [24,27,28].

In this study, we present a previously undescribed EF-hand protein that is highly conserved in filamentous cyanobacteria. Based on demonstrations of Ca^{2+}-binding, we called the protein "Ca^{2+} sensor EF-hand" (CSE). Over-expression of CSE in *Anabaena* was shown to affect photosynthetic electron transfer routes, leading to improper formation of light-harvesting PBS complexes and disrupting oligomerisation of the photosystems, which is essential for the connection of PBS to the reaction centres for functional electron transfer.

2. Materials and Methods

2.1 Growth conditions and treatments of Anabaena cultures

Anabaena cultures were grown in $BG11_AC$ medium, which is $BG11_0$ supplemented with 10 mM $NaHCO_3$ ($BG11_0C$) [29] and 6 mM NH_4Cl. $CoCl_2 \cdot 6\ H_2O$ was replaced with $Co(NO_3)_2 \cdot 6\ H_2O$, and medium contained 10 mM TES-KOH, pH 8.0, and. Cultures were grown under constant illumination of 50 µmol photons m^{-2} s^{-1} with gentle agitation (120 rpm) at 30°C in air enriched with 3% CO_2. Liquid media for the the over-expressor strain (see below) included 1 µg ml^{-1} erythromycin. Fresh *Anabaena* cultures were started at an optical density (OD_{750}) of 0.1 in $BG11_AC$

or $BG11_0C$ (for nitrogen-fixing conditions) as indicated. Total proteins and dry weight of cultures were measured as described in [30]. Chlorophyll *a* absorption (OD_{665}) was measured in cultures suspended in 90% methanol, and pigment absorption spectra of cell suspensions were measured in whole-cell cultures with a Genesys 10S UV-Vis Spectrophotometer (Thermo Scientific) and a Shimadzu UV-1800 UV spectrophotometer (Berner), respectively.

2.2 Generation of Anabaena asr1131 over-expression and knockout strains

The *asr1131* gene sequence flanked by its native promoter and terminator sequences was amplified by PCR with the primers EF1 and EF4 (Table S1), cloned into the vector pST Blue and then into the *RSF1010*-based low-copy number plasmid *pRL1342* [31] using the *Xho*I and *Kpn*I restriction enzyme sites. This resulted in the overexpression plasmid *pBG2089* (see Figure S1), which was used for triparental conjugation of *Anabaena* wild-type (WT) as described in [32]. Transformants with overexpression of *asr1131* were selected on solid growth media containing 10 µg ml^{-1} erythromycin. Mutant strains of *Anabaena* lacking the *asr1131* gene were generated by replacing *asr1131* as well as non-coding DNA of 93 and 266 bp up- and downstream, respectively, with a neomycin/kanamycin-resistance cassette, and introducing the obtained plasmid via triparental conjugation. The *asr1131* knockout mutant strains were not thoroughly investigated in the current study (see Discussion).

2.3 Overexpression of recombinant protein

The *asr1131* coding sequence was amplified by polymerase chain reaction (PCR) using the oligonucleotides *cse*-NdeI-S and *cse*-EcoRI-AS (Table S1), and cloned into the pET-28a(+) vector (Novagen) for introducing a poly-His affinity tag. Recombinant His-tagged asr1131 protein was overexpressed in *Escherichia coli* (*E. coli*) BL21 cells grown at 37°C in Luria-Bertani (LB) medium supplemented with 50 µg ml^{-1} kanamycin by induction with 100 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and purified on a non-commercial Ni-NTA based affinity chromatography system.

2.4 $^{45}Ca^{2+}$ overlay assay

Ca^{2+}-binding capacity was tested in a radioactive $^{45}Ca^{2+}$ overlay assay according to the method described in [33]. 20 – 80 µg proteins were spotted onto a methanol-activated PVDF membrane. The membrane was soaked three times in a Ca^{2+} washing buffer containing 60 mM KCl, 5 mM $MgCl_2$ and 60 mM imidazole-HCl (pH 6.8) for 20 min at room temperature under gentle agitation. Afterwards, the membrane was incubated for 10 min at room temperature in the same buffer supplemented with 0.1 mM $CaCl_2$ and 0.1 µM $^{45}CaCl_2$ (13.90 mCi mg^{-1}; Perkin Elmer), followed by 5 min rinsing with 50% ethanol. Subsequently, the PVDF membrane was completely dried between Whatman No. 1 filter paper and exposed overnight on a phospho-imaging screen. $^{45}Ca^{2+}$ signals

were detected with a FUJI FLA-3000 (FUJIFILM). Proteins were stained with Coomassie Brilliant Blue after the assay.

2.5 Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed in Tris-HCl buffer (pH 7.9), using a VP-ITC microcalorimeter (MicroCal) as described [34] after extensive dialysis of the purified recombinant protein in Chelex 100 (Sigma) to remove contaminating divalent cations. Calorimetric data were evaluated using MicroCal software (OriginLab) and fitted into one-site and two sequential binding sites models, for calculation of the binding thermodynamics [35]. ITC runs were repeated at least three times with two different batches of purified recombinant protein. From fitted curves, the association constants (K_a) were generated and inverted to determine the dissociation constants (K_d).

2.6 Size exclusion chromatography and multi angle light-scattering

Analytical size exclusion chromatography (SEC) using an ÄKTA chromatography system (GE Healthcare Life Sciences) fitted with a Superose 6 Increase 10/300 GL geometric column of 24 ml bed volume (GE Healthcare) was coupled to multi angle light-scattering (MALS) setup comprising a miniDawn Treos system (Wyatt Technology) and an Optilab T-rEX refractometer (Wyatt Technology). SEC MALS experiments were performed at room temperature using a flowrate of 0.5 ml/min, after equilibration of the column with 20 mM Tris-HCl buffer (pH 7.9), as indicated. The elution volume was plotted against the UV signal and the molecular mass was derived from the light scattering data. Data analysis and molecular weight calculations were done using ASTRA software (Wyatt) [36].

2.7 Circular dichroism spectra

Circular dichroism (CD) spectra of purified recombinant protein in 20 mM Tris-HCl buffer (pH 7.9) were recorded at 20°C from 195 to 250 nm using a J-810 spectropolarimeter (JASCO). Thermally-induced protein denaturation was determined by CD spectroscopy at 212 nm, recorded between 20 - 95°C. CD spectra of thermally-denatured proteins were recorded at 95°C, and protein refolding was recorded in samples subsequently cooled to 20°C [37].

2.8 Microscopy techniques

For the determination of the heterocyst frequency, cultures grown for two days in 3% CO_2 in $BG11_AC$ or $BG11_0C$ (for nitrogen-fixing conditions) and cells were stained for 5 min using 0.5% Alcian Blue stain in 50% ethanol and washed three times with the respective growth media. Proheterocysts and heterocysts were counted from ×400 magnification micrographs taken with a

Wetzlar light microscope (Leitz). 1000-2000 cells were counted for each treatment, and the heterocyst frequency calculated as a percentage of total cells counted.

2.9 Nitrogenase activity measurements

Nitrogenase activity was determined using the acetylene reduction assay described by [38]. 5 ml liquid cultures grown in 3% CO_2 in $BG11_AC$ or $BG11_0C$ (for nitrogen-fixing conditions) were flushed with argon for 20 min and incubated in 23 ml vials with 10% acetylene in the headspace for 20 h, under 50 µmol photons m^{-2} s^{-1} light at 30°C with gentle agitation (120 rpm). 20 µl samples of headspace were analysed for ethylene content using a gas chromatograph (GC, Perkin Elmer Clarus® 580) with a Carboxen 1010 PLOT 30 m x 0.53 mm capillary column and a flame ionisation detector (FID) using argon as a carrier gas. 1% ethylene was measured for calibration. The enzyme activity was calculated from the peak area and normalised to protein content.

2.10 RNAseq transcriptomics

Total RNA was isolated as described in [30] from four biological replicates of WT and *asr1131* over-expressor cultures grown for two days in $BG11_AC$ medium in 3% CO_2. RNA samples were submitted to the Beijing Genomics Institute (Shenzhen, China) for single-ended library preparation and sequencing of RNA libraries using Illumina-HiSeq2500/4000. RNAseq reads were aligned with the Strand NGS 2.7 software (Agilent, USA) using the reference genome and annotations of *Nostoc* sp. PCC 7120, downloaded from Ensembl (EBI). The DESeq R package was used for normalisation and quantification of the aligned reads. Significantly differentially expressed genes were identified using a 2-way ANOVA test with Benjamini-Hochberg p-value correction for the calculation of the false discovery rate (FDR).

2.11 Photosynthetic fluorescence analysis

Cultures were grown for two days in 3% CO_2 in $BG11_AC$ and adjusted to a chlorophyll *a* concentration of 7.5 µg ml^{-1} prior to measurements of low temperature (77K) and room temperature fluorescence emission spectra, chlorophyll *a* fluorescence and P700 absorbance, F_0 rise, state transitions, light curves, single flash-induced fluorescence decay, and P700 oxidoreduction according to [39,40]. For fluorescence emission spectra, 5 µM Eosin Y was used as an internal standard at 536 nm for excitation with 440 nm light generated with a monochromator. Chlorophyll fluorescence and P700 absorbance were measured independently using 400 ms saturating pulses, with samples for all P700 measurements adjusted to 15 µg ml^{-1} chlorophyll *a*. The acceptor side limitation of P700 (Y(NA)) was calculated as $(P_m-P_m')/P_m$.

Oxygen evolution in light and oxygen consumption in darkness were measured with a Clark-type oxygen electrode (DW1, Hansatech) with constant stirring at 30°C. Cultures were adjusted to a

chlorophyll concentration of 7.5 µg ml^{-1}. Oxygen concentrations were recorded in darkness for 5 min before subsequent illumination with 400 µmol photons m^{-2} s^{-1}.

2.12 Photosynthetic protein complex analysis and proteomics

Thylakoid membrane protein fractions were isolated from 2 day-old cultures according to [40]. 75 µg thylakoids were solubilised with 1.5% *n*-dodecyl-β-D-maltoside (DM) and separated by large pore blue native polyacrylamide gel electrophoresis (lpBN-PAGE) using an acrylamide concentration gradient of 3.5-12.5%. After lpBN-PAGE separation, gel strips were cut, and proteins denatured in Laemmli buffer containing 5% β-mercaptoethanol and 6 M urea at room temperature for 30 min. After solubilisation, the strips were laid onto a 12% SDS polyacrylamide gel containing 6 M urea, and proteins were separated in the second dimension and subsequently stained with Sypro Ruby Protein Gel Stain. For identification of proteins by mass spectrometry, protein spots were excised from SDS-PAGES gels and proteins subjected to in-gel trypsin digestion, as previously described [42]. Peptides were identified by nanoscale liquid chromatography/electrospray ionization tandem mass spectrometry (nLC/ESI-MS/MS) using a Q-Exactive instrument (Thermo Scientific). The MS/MS spectra were analysed using *Nostoc* sp. PCC 7120 annotations (GCA_000009705.1 downloaded from Cyanobase; http://genome.microbedb.jp/cyanobase/) using Proteome Discoverer v.2.2 (Thermo Scientific) as previously described [42].

2.13 Bioinformatics methods

The tertiary structures of translated gene coding sequences were predicted using the I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/). Asr1131 homologs obtained from Basic Local Alignment Search Tool (BLAST) searches were aligned by Muscle in MEGA6. Gene sequences, identifications and descriptions were collected from CyanoBase (Kazusa Genome Resources; genome.microbedb.jp/cyanobase), KEGG (www.genome.jp/kegg/), UniProt (uniprot.org) and the National Center for Biotechnology Information (NCBI; ncbi.nlm.nih.gov) databases.

3. Results

3.1 A small Ca^{2+}-binding protein is highly conserved in filamentous cyanobacteria

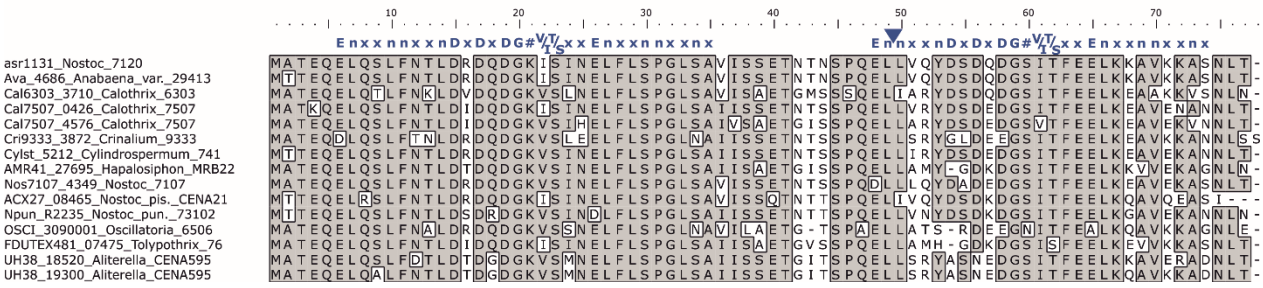
In silico searches for Ca^{2+}-binding domains in cyanobacterial genome databases identified a gene putatively encoding a protein with two Ca^{2+}-binding EF-hand domains, which was previously not described in published literature. 15 homologs of the putative EF-hand protein were encoded in 13 cyanobacterial species, with two paralogs each occurring in *Calothrix* sp. PCC 7507 and *Aliterella atlantica* CENA595. The gene was found exclusively in filamentous cyanobacteria, with one

exception being the unicellular species *Aliterella atlantica* CENA595, which is closely related to filamentous cyanobacteria [43,44]. Two of the filamentous species encoding the putative EF-hand protein were non-heterocystous (*Crinalium epipsammum* PCC 9333 and *Oscillatoria* sp. PCC 6506), although *Oscillatoria* sp. PCC 6506 is capable of fixing atmospheric N [45].

Alignment of amino acid sequences translated from all orthologues of the putative EF-hand protein (Figure 1A) demonstrated strong conservation of the N-terminal EF-hand domain, which typically comprises 12 amino acids that form a Ca^{2+}-binding loop flanked on each side by nine amino acids forming α-helices [46]. The C-terminal EF-hand encoded in orthologues from *Crinalium*, *Hapalosiphon*, *Oscillatoria*, *Tolypothrix* and *Aliterella* displayed variation in the first five positions of the loop (positions 54 – 58 in Figure 1A), of which positions 1, 3 and 5 are known to be directly involved in binding of Ca^{2+} [46,47]. Orthologue sequences retrieved from genomes of several strains of *Anabaena, Nostoc*, *Calothrix* and *Cylindrospermum* possessed the conserved positions required for binding Ca^{2+} in both EF-hands. However, the α-helix preceding the second Ca^{2+}-binding loop lacked three amino acids (between position 49 – 50 in Figure 1B), and the canonical hydrophobic residue preceding the first Ca^{2+} ligand was substituted with a tyrosine (see position 53 in Figure 1B).

In the filamentous cyanobacterium *Nostoc* sp. PCC 7120 (*Anabaena*) the putative EF-hand protein is encoded by *asr1131* and has a theoretical atomic mass of 8.5 kDa. Threading analysis to predict the tertiary structure of translated *asr1131* showed strongest resemblance with the helix-loop-helix structures of Ca^{2+}-binding domains in calmodulin and troponin C (Figure 1B).

A



B

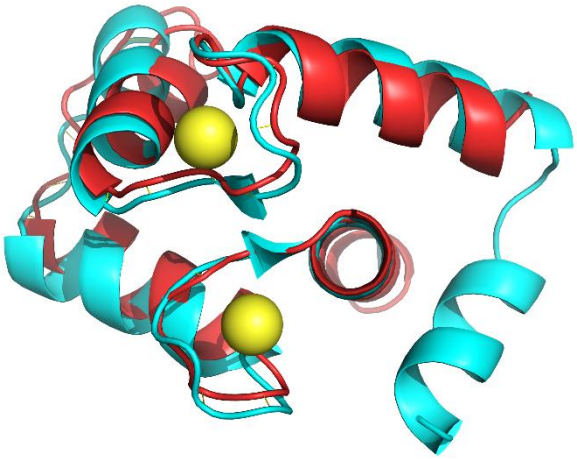


Figure 1 Multiple amino acid sequence alignment and predicted structure of putative EF-hand protein conserved in filamentous cyanobacteria

A. Multiple sequence alignment of *asr1131* and other homologues from cyanobacteria. Identical residues are boxed and shaded. Canonical EF-hand sequence (blue letters) is aligned above the two EF-hand sequences of CSE, where specific amino acids in are represented by their single letter code, n is a hydrophobic residue, x is any residue, # is the fourth Ca^{2+} ligand (http://www.structbio.vanderbilt.edu/cabp_database/seq/indiv.aa/index.html). The blue triangle indicates a missing "xxn" motif from the second EF-hand
B. Translated *asr1131* amino acid sequence (red) threaded through the crystal structure of the calcium-binding domain of mammalian troponin C (blue; PDB 2TN4) containing two bound Ca^{2+} (yellow);

3.2 In vitro biochemical properties of recombinant asr1131 *protein*

The *asr1131* gene was cloned and over-expressed with an N-terminal His-tag. Purified recombinant protein was applied to size exclusion chromatography coupled with multiangle light scattering (SEC-MALS) for the determination of the oligomerisation state of recombinant asr1131 protein in the presence or absence of Ca^{2+}. In Ca^{2+}-Tris buffer, His-tagged asr1131 eluted as a monomer with an apparent mass of 9.65 kDa (Figure 2A). In Mg^{2+}-Tris buffer lacking Ca^{2+}, a minor shift in the elution profile of asr1131 was detected, corresponding to a molecular mass of 10.27 kDa, indicating a conformational change in the protein induced by Ca^{2+} ions (Figure 2A). The eluted peaks were collected and subjected to SDS-PAGE to confirm the presence of asr1131 protein in each fraction (Figure 2B). SEC-MALS analysis showed the protein to be monomeric in solution.

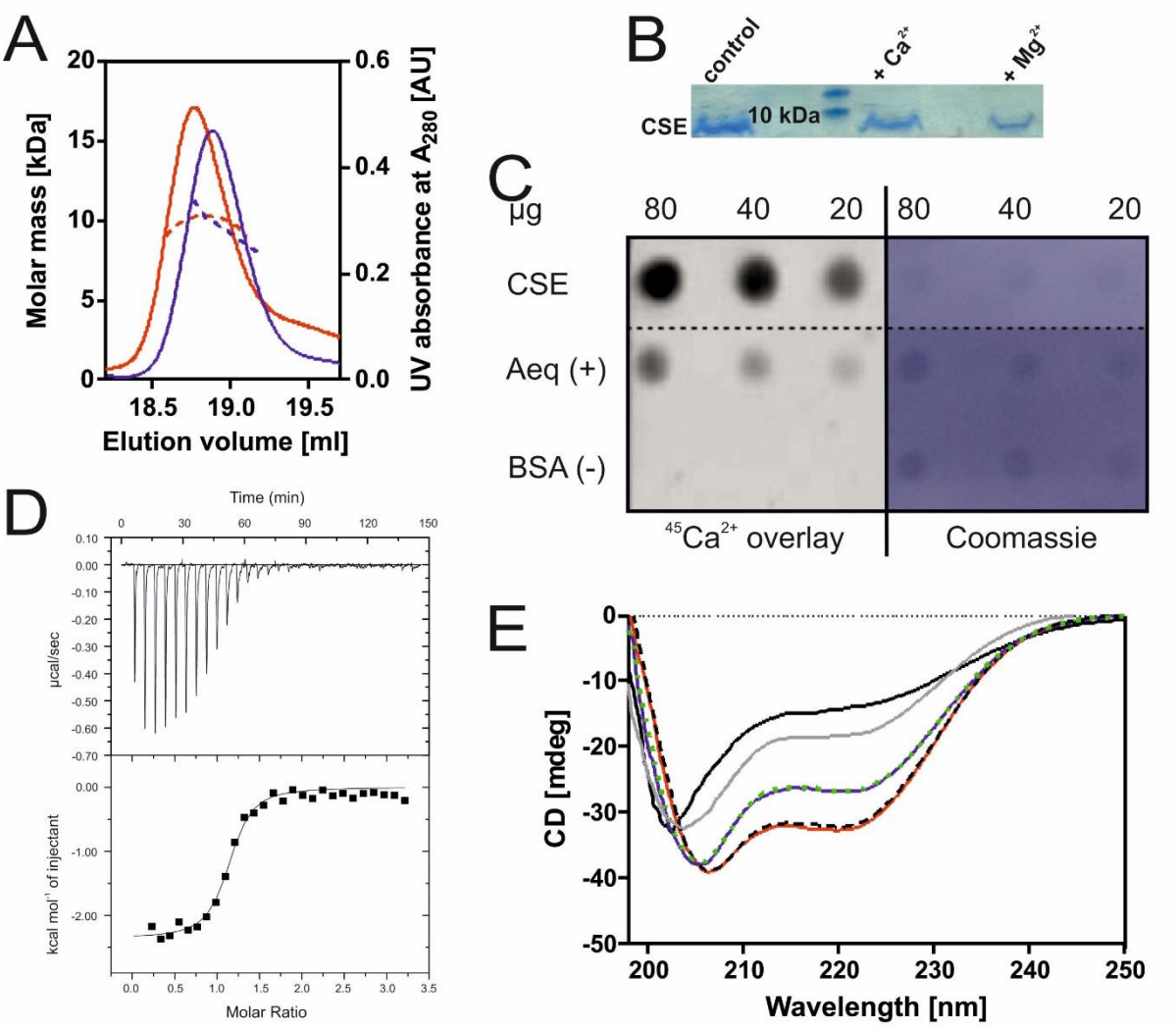


Figure 2 Structural analysis and Ca^{2+}-binding properties of purified recombinant CSE protein

(A) Size exclusion chromatography separation coupled to multiangle light scattering (SEC-MALS) used to detect the monomeric state of CSE in the presence of Ca^{2+} (blue line) and Mg^{2+} (red line), the dotted lines represent the distribution of molar mass across the respective peaks. **(B)** SDS-PAGE of collected protein peaks eluted from SEC-MALS; **(C)** $^{45}Ca^{2+}$ overlay assay of purified recombinant CSE protein. Aequorin (Aeq) and BSA were used as positive and negative controls, respectively. Image represents the results from three technical replicates. Proteins shown are from a single membrane that were treated together, and the image has been cropped to remove irrelevant proteins. **(D)** Isothermal titration calorimetry (ITC) metal-binding assay of recombinant His-tagged CSE (66 µM) in 1 mM $CaCl_2$ in 20 mM Tris buffer. **(E)** Circular dichroism (CD) spectra at 20°C in different Tris buffers; without additions (black solid line); with NaCl (grey line); with NaCl and Mg^{2+} (blue line); with Mg^{2+} (green dashed line); with NaCl and Ca^{2+} (red line), with Ca^{2+} (black dashed line)

Three dilutions of purified recombinant His-tagged asr1131 were spotted onto a PVDF membrane that was subsequently overlaid with radioactive $^{45}Ca^{2+}$. The autoradiogram showed efficient binding of Ca^{2+} to recombinant asr1131 *in vitro* (Figure 2C, top row). The weakest $^{45}Ca^{2+}$ signal was

obtained with the lowest protein amount (20 µg protein), and the signal intensity increased with each two-fold increase in protein amount. The $^{45}Ca^{2+}$ signal obtained from recombinant asr1131 was even stronger than from the Ca^{2+}-sensing photoprotein aequorin (AEQ), which was used as a positive control. Bovine serum albumin (BSA) used as a negative control showed no non-specific binding of $^{45}Ca^{2+}$.

Table 1 Thermodynamic parameters for Ca^{2+}-binding to CSE protein

The raw isothermal titration calorimetry (ITC) data were fitted using a one-site binding model for monomeric CSE. [a] The dissociation constant (K_d) values correspond to the mean of the independent experiments ± SD. All titrations were performed in 20 mM Tris-HCl based buffer (pH 7.9). [b] number of calculated binding sites [c] enthalpy [d] entropy [e] reversible heat [f] Gibbs energy

Titrant/CSE protein	**One-site binding model**					
	Average K_D (µM) [a]	**n** [b]	**ΔH (*kcal mol^{-1}*)** [c]	**ΔS (*cal mol^{-1} K^{-1}*)** [d]	**TΔS (*kcal mol^{-1}*)** [e]	**ΔG (*kcal mol^{-1}*)** [f]
Ca^{2+} (in Tris buffer)	1.4 ± 1.0	0.9	-2.3 ± 0.3	18.2 ± 3.7	5.3 ± 1.1	-7.6 ± 0.9
Mg^{2+} (in Tris buffer)	35.2 ± 2.1	4.1	0.6 ± 0.2	22.2 ± 1.2	6.5 ± 0.3	-5.9 ± 0.2
Ca^{2+} (in presence of 100 mM NaCl)	4.7 ± 2.9	0.8	-1.6 ± 0.3	18.0 ± 3.1	5.2 ± 0.9	-6.9 ± 0.6
Mg^{2+} (in presence of 100 mM NaCl)	No binding					
Ca^{2+} (in presence of 300 µM Mg^{2+})	3.5 ± 3.3	1.1	-3.4 ± 0.5	15.1 ± 3.2	4.4 ± 0.9	-7.9 ± 0.6
Mg^{2+} (in presence of 150 µM Ca^{2+})	No binding					
Ca^{2+} (in presence of 300 µM Mg^{2+} and 100 mM NaCl)	9.3 ± 1.4	1.2	-2.1 ± 0.3	16.2 ± 1.7	4.7 ± 0.5	-6.8 ± 0.4
Ca^{2+} (in presence of 2 mM Mg^{2+} and 100 mM NaCl)	14.4 ± 7.8	1.4	-2.1 ± 0.3	16.2 ± 1.7	4.7 ± 0.5	-6.8 ± 0.4
Mg^{2+} (in presence of 150 µM Ca^{2+} and 100 mM NaCl)	No binding					

Isothermal titration calorimetry (ITC) using His-tagged recombinant asr1131 protein was performed to define its Ca^{2+}-binding stoichiometry and affinity. Raw ITC data were fitted into a one-binding site model to determine the stoichiometry of bound ligands (n), and the average dissociation constant (K_d) for all available binding sites (Table 1, Figure S2). It was reported previously that

NaCl can inhibit Ca^{2+}-binding of the *Anabaena* Ca^{2+}-binding protein CcbP [48], therefore the Ca^{2+}- and Mg^{2+}-binding properties of asr1131 were examined in the absence/presence of NaCl. In absence of NaCl, titration of asr1131 with ~5 µM $CaCl_2$ per injection yielded exothermic calorimetric signals that saturated at a molar ratio of 1.11 (Figure 2D). Curve-fitting yielded a K_d of 1.4 ± 0.98 µM at a molar ratio of n = 0.94 (Table 1). These results corresponded to the binding of a single Ca^{2+} ion per asr1131 protein molecule. In contrast, the Mg^{2+}-binding events showed an endothermic profile with a K_d value of 35.2 ± 2.1 µM, indicating very weak binding of Mg^{2+} (Table 1, Figure S2A), and supporting the specificity of asr1131 for Ca^{2+}. Similar trends were exhibited in Tris buffer containing NaCl (Figure S2B); however, the binding enthalpy of Ca^{2+} was reduced (compare Figure 2D with S2B) and K_d values of 4.7 ± 2.9 µM were substantially higher than K_d in the absence of NaCl. Mg^{2+}-binding was completely abolished by the addition of NaCl (Figure S2). Intracellular concentrations of free Mg^{2+} are much higher than Ca^{2+} *in vivo*, therefore Ca^{2+}-binding by recombinant asr1131 under high Mg^{2+} were assessed. In 300 µM Mg^{2+}, strong exothermic signals for each Ca^{2+} injection showed that Ca^{2+} bound to the asr1131 protein (Figure S2C), with a K_d value of 3.5 ± 3.3 µM and a molar ratio of n = 1.1 (Table 1) corresponding to binding of a single Ca^{2+} per protein molecule even in the presence of high Mg^{2+}. Similar results were obtained in the presence of 100 mM NaCl, although with lower binding affinity of Ca^{2+} (K_d = 9.3 ± 1.4) (Table 1, Figure S2D). By contrast, the binding of 3 mM Mg^{2+} to asr1131 under 150 µM Ca^{2+} was completely suppressed, and Mg^{2+} was also unable to compete with Ca^{2+} for protein-binding in Tris buffer supplemented with 100 mM NaCl and 150 µM $CaCl_2$ (Figure S2 and Table 1). Moreover, Ca^{2+} was able to induce strong exothermic binding signals and a K_d of 14.4 µM even in presence of 2 mM Mg^{2+} (Table 1, Figure S2E). Because the primary asr1131 sequence predicted two Ca^{2+}-binding sites, the raw ITC data was also fitted to a model with two sequential binding sites, resulting in K_d values of 5.9 ± 7.5 and 56.3 ± 50.1 for the N-terminal and C-terminal binding site, respectively. Based on the ITC experiments, we concluded that under physiological conditions, recombinant asr1131 specifically binds one Ca^{2+} ion to the N-terminal binding site with high affinity, while the C-terminal site may have very low affinity to Ca^{2+} ions.

To identify the effects of Ca^{2+} on the protein secondary structure, the CD spectra of recombinant asr1131 were recorded in the presence or absence of Ca^{2+} and Mg^{2+}. CD spectra exhibited typical α-helical profiles with two characteristic minima at 207 nm and 222 nm (Figure 2E). Ca^{2+} and Mg^{2+} caused a shift in the CD spectra towards increased α-helix peaks, indicating that both $Ca^{2+}$and Mg^{2+} induced conformational changes in the protein, with Ca^{2+} inducing greater peak shifts than Mg^{2+}. The melting curve of recombinant asr1131 protein from 20-95°C in the presence and absence of Ca^{2+} showed that a large proportion of the protein remained folded at 95°C (Figure S3A).Upon subsequent cooling to 20°C, the unfolded fraction of the protein regained the correct folding (Figure S3B), indicating that the asr1131 protein exhibits exquisite stability. Based on the

Ca^{2+}-binding activity and secondary structure determined here, we named the asr1131 protein "Ca^{2+} sensor EF hand" (CSE).

3.3 Overexpression of cse in Anabaena affects culture growth and pigment contents, but not heterocyst abundance or function

To explore the function of CSE *in vivo*, an *Anabaena* strain with upregulated expression of *cse* was engineered by transformation of wild-type *Anabaena* with the low copy number plasmid *pBG2089* (Figure S1) containing the *asr1131* gene under the control of its native promoter. This strain was found to exhibit two-fold higher expression of *cse* than WT *Anabaena* (Table 3). The *cse* over-expressor (*cse*-ox) strain was characterised in relation to WT *Anabaena*. Evaluation of culture growth rates over three days in CO_2-enriched air (3% CO_2), in $BG11_AC$ (N-replete) or $BG11_0C$ (N-fixing) media revealed about 30-50% slower growth of *cse*-ox in both media, compared to the WT (Figure 3A). Room temperature whole cell absorption spectra normalised to 750 nm showed a significant reduction in the PBS peak at 635 nm in *cse*-ox in both N-fixing and N-replete conditions (Figure 3B). The over-expressor also had a significantly lower chlorophyll peak at 685 nm, compared to the WT, when grown under N-fixing growth conditions. The carotenoid peak at 495 nm in the *cse*-ox was not significantly different from WT.

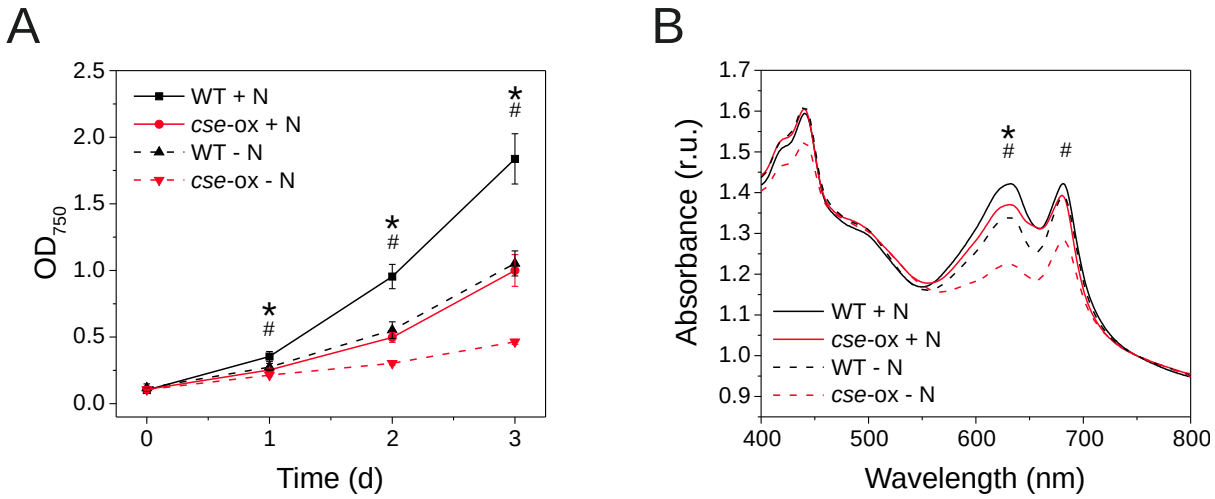


Figure 3 Growth phenotype induced by *cse* overexpression in *Anabaena* grown in nitrogen-replete and nitrogen-fixing conditions

(A) Growth of *Anabaena* wild-type (WT; black) and *cse* overexpressor (*cse*-ox; red) strains monitored by absorbance at 750 nm. Cultures were grown in 3% CO_2 in $BG11_AC$ medium (+N; solid lines) or $BG11_0C$ (- N; dashed lines); **(B)** Absorption spectra of WT (black) and *cse*-ox (red) grown for two days in 3% CO_2 in $BG11_AC$ medium (+N; solid lines) or $BG11_0C$ (- N; dashed lines) normalised to absorbance at 750 nm. Peaks correspond to chlorophyll (440 nm and 680 nm), carotenoids (500 nm) and phycobilisomes (630 nm). Data points represent mean values from four biological replicates, error bars in **(A)** show standard deviations. Significant differences between WT and mutant samples are indicated with asterisks (+ N) and hashes (- N), respectively, (*t*-test $P<0.05$)

Comparisons of the protein and chlorophyll contents, the dry weight, heterocyst frequency and nitrogenase activity showed no significant differences between WT and *cse*-ox in N-replete conditions (Table 2). However, *cse*-ox cells grown for 2 days in N-fixing conditions had about 30% less chlorophyll and were around 12% heavier than WT cells. Heterocyst frequency and nitrogenase activity did not differ significantly between the two strains.

Table 2 Physiological parameters in N-replete and N-fixing conditions

2 day-old *Anabaena* WT and *cse* overexpressor (*cse*-ox) cultures were grown in 3% CO_2 in $BG11_0C$ medium either supplemented with NH_4^+ (+N) or lacking a source of combined N (-N). Errors indicate standard deviations that were calculated from three biological replicates, significant differences between WT and mutant samples are indicated with asterisks (*t*-test $P<0.05$)

Strain and condition	**Protein content (µg/ml) per OD_{750} = 1.0**	**chl content (µg/ml) per OD_{750} = 1.0**	**Dry weight (mg/ml) per OD_{750} = 1.0**	**Heterocyst frequency (% of total cells counted)**	**Nitrogenase activity (µmol h^{-1} mg $proteins^{-1}$)**
WT + N	273.1 ± 23.6	9.1 ± 0.3	0.46 ± 0.01	1.3 ± 0.5	n.d.
cse-ox + N	304.6 ± 11.3	9.1 ± 0.3	0.49 ± 0.02 *	1.6 ± 0.3	n.d.
WT - N	218.5 ± 9.0	9.3 ± 0.2	0.44 ± 0.01	9.4 ± 1.0	2.49 ± 0.09
cse-ox - N	208.5 ± 10.1	6.3 ± 0.2 *	0.50 ± 0.01 *	9.3 ± 1.1	2.99 ± 0.37

3.4 Expression of genes encoding photosynthetic light-harvesting proteins is differentially regulated by cse overexpression

The transcriptomes of *Anabaena* WT and the *cse-ox* strain grown in $BG11_0C$ + NH_4^+ media in 3% CO_2 were analysed after two days of growth in fresh medium. Significant differences were identified in expression of genes encoding subunits of the light-harvesting PBS complexes in *cse-ox*, compared to WT (Table 3). Members of the *cpc* gene cluster (*alr0528* – *alr0536*) encoding PC rods, and the entire *pec* cluster (*alr0523-alr0527*), encoding PEC rod caps [24] were strongly downregulated. Three genes encoding OCP-like N-terminal domain proteins (*all1123*, *all3221*, *alr4783*) were downregulated in the *cse*-ox, along with photosynthetic complex subunits *psaB* and *ndhD1* and several genes from the porphyrin pathway including *hemK*, *por*, *chlL* and *chlN*. Conversely *hemF2*, which is also involved in chlorophyll and porphyrin biosynthesis, was strongly upregulated in *cse*-ox compared to WT. Several genes involved in Ca^{2+}-related processes were also differentially expressed in the over-expressor (see Table 3).

Table 3 Transcription changes in the *cse*-ox mutant

Differential expression of genes in *cse*-ox, in comparison to WT, grown for 2 days in $BG11_{A}C$ medium in 3% CO_2. Genes with log2 fold change (FC) values ≥0.9 (upregulated) or ≤-0.9 (downregulated) are shown shaded red and blue, respectively. False discovery rates (FDR) were calculated using the Benjamini-Hochberg method. For the gene cluster marked with an asterisk, the average log2FC and largest FDR value are shown

Accession	**Gene ID**	**Description**	**log2 FC *cse*-ox/ WT**	**FDR**
asr1131	*cse*	calcium sensor EF-hand	1.2	1.30E-06
Phycobilisomes				
alr0528	*cpcB*	phycocyanin beta chain	-0.9	7.93E-06
alr0529	*cpcA*	phycocyanin alpha chain	-0.9	5.02E-06
alr0530	*cpcC*	phycocyanin-associated rod linker protein	-2.2	2.96E-07
asr0531	*cpcD*	rod-capping linker polypeptide	-1.8	6.09E-06
alr0532	*cpcE*	phycocyanobilin lyase alpha subunit	-1.7	9.62E-07
alr0533	*cpcF*	phycocyanobilin lyase beta subunit	-1.7	4.80E-06
alr0534	*cpcG1*	phycobilisome rod-core linker protein	-1.2	1.75E-05
alr0535	*cpcG2*	phycobilisome rod-core linker protein	-1.1	3.09E-05
alr0536	*cpcG3*	phycobilisome rod-core linker protein	-1.1	1.47E-05
alr0523	*pecB*	phycoerythrocyanin beta chain	-2.4	3.37E-06
alr0524	*pecA*	phycoerythrocyanin alpha chain	-2.4	9.95E-073
alr0525	*pecC*	phycoerythrocyanin-associated rod linker protein	-2.6	5.01E-07
alr0526	*pecE*	bilin biosynthesis protein	-2.8	5.66E-07
alr0527	*pecF*	bilin biosynthesis protein	-2.8	2.46E-07
all1123		OCP NTD-homolog	-1.7	7.61E-05
all3221		OCP NTD-homolog	-1.0	2.03E-04
alr4783		OCP NTD-homolog	-1.7	1.19E-05
Chlorophyll and porphyrin biosynthesis				
alr0115	*hemK*	protoporphyrinogen oxidase	-1.4	9.29E-07

all1357	*hemF2*	coproporphyrinogen III oxidase	1.5	1.35E-05
all1743	*por*	protochlorophyllide oxido-reductase	-1.1	1.92E-06
alr3125		heme oxygenase	-1.1	4.66E-04
alr3751		ferrochelatase	-1.3	1.26E-07
all5076	*chlN*	protochlorophyllide reductase subunit	-1.0	1.85E-05
all5078	*chlL*	protochlorophyllide reductase iron-sulfur ATP-binding protein	-1.3	1.21E-05
Photosynthesis				
alr0348	*ndhD*	NADH dehydrogenase type I subunit 4	-1.5	1.57E-04
asl3190		similar to photosystem I subunit IX (*psaJ*)	-1.3	1.37E-04
alr5155	*psaB*	photosystem I core protein A2	-1.0	1.00E-04
Nitrogen fixation				
all0571		cyanophycinase	-1.6	1.02E-04
asr0773		Mo-dependent nitrogenase-like protein	1.6	1.08E-04
alr0874	*nifH2*	nitrogenase reductase	-1.2	0.003975
all1440	*nifK*	nitrogenase molybdenum-iron protein beta chain	-1.1	0.003639
all1455	*nifH1*	nitrogenase iron protein	-1.9	1.09E-04
Other genes of interest				
alr1381	*prcA*	trypsin; Ca^{2+}-dependent protease which degrades phycobiliproteins *in vitro* (Maldener *et al.*, 1991)	1.3	5.08E-05
alr3199		hemerythrin DNase with Ca^{2+} and Fe^{2+}-dependent nickase activity (Padmaja *et al.*, 2011)	-3.1	2.32E-05
alr3731	*prpA*	protein serine-threonine phosphatase	1.1	7.01E-04
**alr0198 – alr0199*		unknown proteins; surface-associated protein; in PC complex; responsive to Ca^{2+} treatments (Walter *et al.*, 2016)	-2.4	4.66E-05

3.5 Photosynthetic pigment-protein complexes and photosynthetic activity are altered in cse-ox

Energetic connectivity within the photosynthetic pigment-protein complexes was studied by fluorescence emission at low temperature (77K) and room temperature in whole cell samples excited with monochromatic light of 440 nm (specific for chlorophylls) or 580 nm (specific for PBS).

Chlorophyll excitation revealed a PSII:PSI ratio of about 1:4 in both the WT and *cse*-ox (Figure 4A). PBS excitation at 77K showed the peak at 646 nm, which corresponds to PC [23], to be severely decreased in *cse*-ox compared to WT, while the APC peak at 663 nm was significantly higher in the *cse*-ox (Figure 4B). At room temperature, the APC peak (660 nm) was also significantly higher in the over-expressor than in WT, while the shoulder for PC (640 nm) was significantly lower in *cse*-ox.

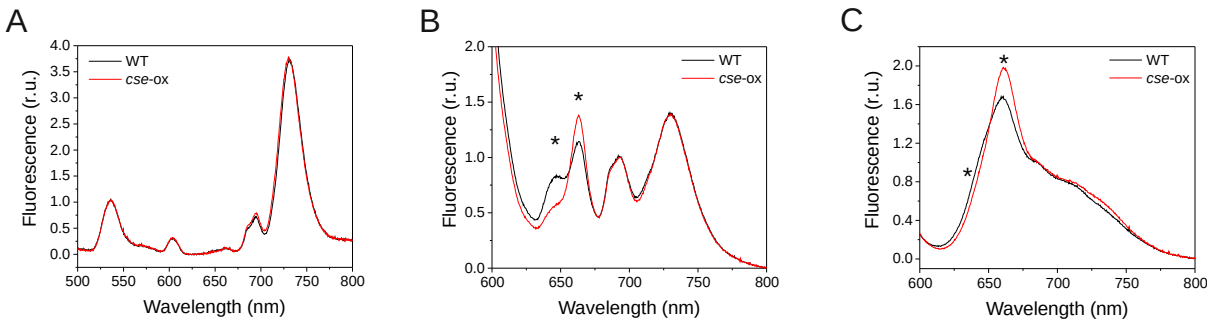


Figure 4 Fluorescence emission spectra of 2 day-old *Anabaena* wild-type (WT) and *cse* overexpression (*cse*-ox) strains grown in $BG11_AC$

(A) Emission at low temperature (77K) after chlorophyll excitation with 440 nm monochromatic light. Spectra are normalised to internal standard Eosin Y (5 µm, peak at 536 nm), emission peaks from PSII and PSI are located at 694 nm and 730 nm, respectively; **(B)** Emission at low temperature (77K) after phycobilisome excitation with 580 nm monochromatic light. Spectra are normalised to the PSII peak at 694 nm. Other peaks show emission from phycocyanin (646 nm), allophycocyanin (663 nm), phycobilisome terminal emitter (687 nm) and PSI (730 nm); **(C)** Emission at room temperature (25°C) after excitation of phycobilisomes with 580 nm monochromatic light. Spectra are normalised to the peak at PSII peak at 685 nm. Other peaks show emission from allophycocyanin (660 nm), phycocyanin (640 nm) and PSI (715 nm). All curves represent mean values from at least three biological replicates. Significant differences between WT and mutant samples are indicated with asterisks (*t*-test $P<0.05$)

Photosynthetic performance was monitored with a Dual-PAM-100 (Walz) on culture samples that were dark-adapted and then illuminated with constant actinic light (50 µmol photons m^{-2} s^{-1}). The *cse*-ox cultures showed higher fluorescence in the absence of actinic light (F_0 = 1.4) in comparison to the WT (F_0 = 1.1), as well as higher steady state fluorescence under actinic light (F_s), but saturating light pulses induced smaller F_m' values relative F_s (Figure 5A). The fluorescence signal monitored during light-to-dark transitions showed the transient rise in fluorescence that was prominent in the WT after removal of actinic light to be significantly diminished in the *cse*-ox strain (Figure 5A inset). F_v/F_m determination in the presence of DCMU showed that PSII photochemical efficiency was around 65% lower in *cse*-ox in comparison to the WT value (Figure 5B), and the Y(II) parameter calculated during step-wise increases in actinic irradiance also showed that the maximum PSII quantum yield in *cse*-ox was significantly lower than in WT in light intensities from 0

to 220 µmol photons m^{-2} s^{-1}, although no significant difference between the strains was observed at higher light intensities (Figure 5C). Analysis of PSII fluorescence decay after a single flash revealed significantly faster decay in *cse*-ox during the initial phase (1.5 – 3.2 ms), and slower decay in the second phase (0.006 – 3 s) (Figure 5D).

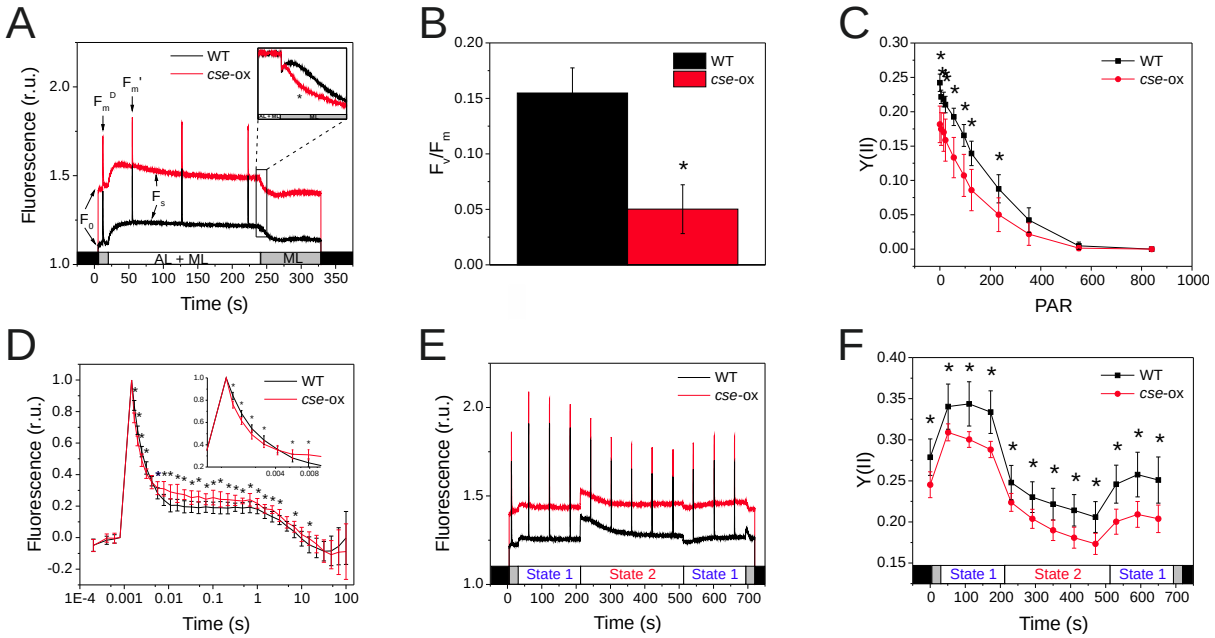


Figure 5 PSII-fluorescence parameters of 2 day-old *Anabaena* wild-type (WT) and *cse* overexpression (*cse*-ox) strains grown in $BG11_AC$

(A) Fluorescence induction during illumination with only measuring light (ML; 5 µmol photons m^{-2} s^{-1}) or with ML + actinic light (AL; 57 µmol photons m^{-2} s^{-1}). Fluorescence level under ML (F_0), maximum fluorescence of open PSII reactions centres (F_m), maximum fluorescence after AL illumination ($F_{m'}$) and steady state fluorescence under AL (F_s) are indicated. Inset: transient rise in fluorescence after switching off AL (F_0 rise normalised to F_s); **(B)** F_v/F_m calculated from fluorescence measurements of dark-adapted cells in the presence of 20 µM DCMU; **(C)** PSII quantum yield (Y(II)) over increasing intensities of actinic light; **(D)** Flash-induced increase and subsequent decay in fluorescence in dark-adapted WT and *cse*-ox samples; **(E)** Chlorophyll fluorescence during state transitions in dark-adapted samples illuminated with blue light (460 nm; State 1) and red light (620 nm; State 2); **(F)** PSII quantum yield (Y(II)) calculated from fluorescence induced by saturating pulses during state transitions measurements. All data represent mean values from four biological replicates, error bars show standard deviations. Significant differences between WT and mutant samples are indicated with asterisks (*t*-test $P<0.05$)

To determine whether the capacity for state transitions was affected in *cse*-ox, chlorophyll fluorescence was measured in dark-adapted samples that were illuminated with actinic blue light to induce state 1, then with actinic red light to induced state 2, and then again with blue light. The results showed equivalent capacity for state transitions in both *cse*-ox and WT, despite a higher basal fluorescence level in the over-expressor, as detected earlier (Figure 5E). Y(II) in *cse*-ox

during state transition measurements was consistently and significantly lower than in the WT in both state 1 and state 2 conditions (Figure 5F).

P700 maximum quantum yield Y(I), and donor and acceptor side limitation (Y(ND) and Y(NA), respectively) were measured by pre-illuminating culture samples with far-red light, followed by illumination with red actinic light and periodic application of saturating pulses. After far red light illumination, *cse*-ox had about 20% higher Y(I) (Figure 6A) and 30% lower Y(ND) (Figure 6B), compared to WT. P_m measurements showed equivalent levels of P700 oxidation in both WT and over-expressor (data not shown). During excitation with red light, *cse*-ox showed significantly lower Y(I) levels and higher Y(NA) in comparison to the WT (Figures 6A and 6C). Y(I) values measured over increasing light intensities showed lower Y(I) in the *cse*-ox between 0 to 220 µmol photons m^{-2} s^{-1}, with values at high light intensities equivalent to WT (Figure 6D). Absorbance of $P700^+$ measured in dark-adapted cells during 5 s strong far red light illumination showed rapid increases P700 oxidation under far red in both genotypes, but also revealed a decrease in $P700^+$ signal in *cse*-ox during far red illumination that was not seen in WT (Figure 6E).

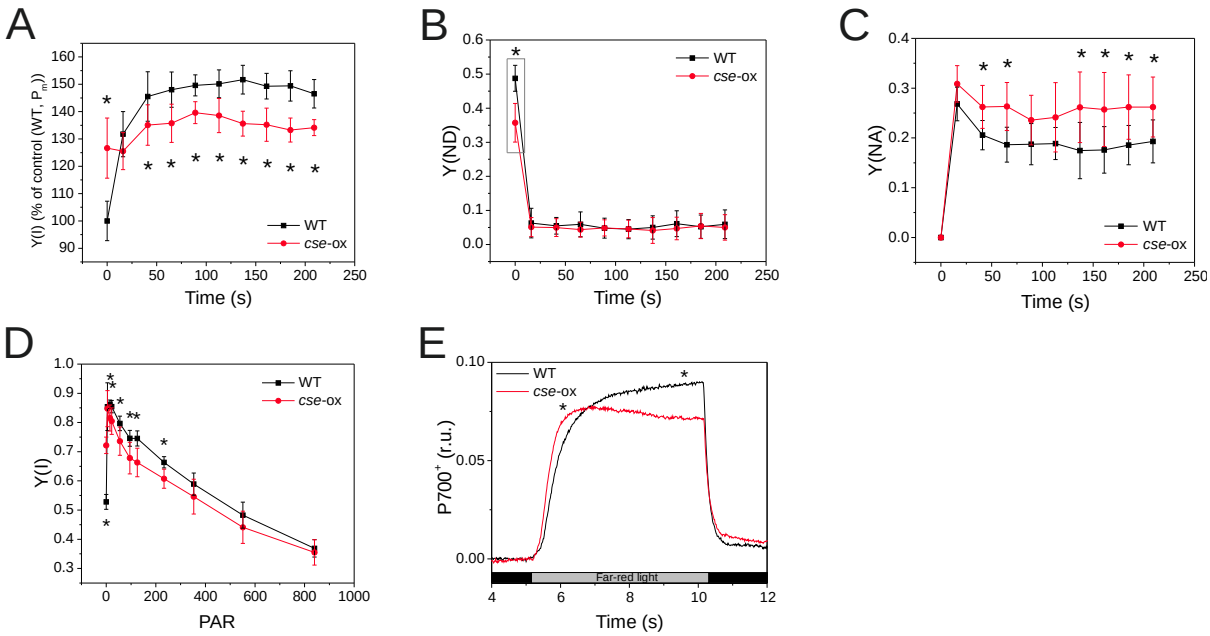


Figure 6 PSI parameters measured in 2 day-old *Anabaena* wild-type (WT) and *cse* over-expression (*cse*-ox) strains grown in $BG11_AC$

(A) P700 maximum quantum yield (Y(I)) in WT and *cse*-ox culture samples during illumination with red actinic light after 10 s illumination with far red light (for determination of maximum $P700^+$). Y(I) values are expressed relative to WT time-point 0, which represents 100% Y(I); **(B)** Donor side limitation of PSI (Y(ND)) and **(C)** acceptor side limitation of PSI (Y(NA)) calculated from $P700^+$ values measured during actinic light illumination in **(A)**; **(D)** P700 maximum quantum yield from (Y(I)) determined over increasing intensities of red actinic light; **(E)** P700 oxidation and reduction during far red illumination, normalised to baseline values only. All data represent mean values from four biological replicates, error bars show standard deviations. Significant differences between WT and mutant samples are indicated with asterisks (*t*-test $P<0.05$)

Oxygen uptake/evolution measured with a Clark-type oxygen electrode revealed about 35% lower respiration rate in *cse*-ox during 5 min dark incubation, compared to the WT control (Figure 7A), while about 20% higher net oxygen evolution rate was recorded in the over-expressor under illumination with 400 µmol photons m^{-2} s^{-1} (Figure 7B).

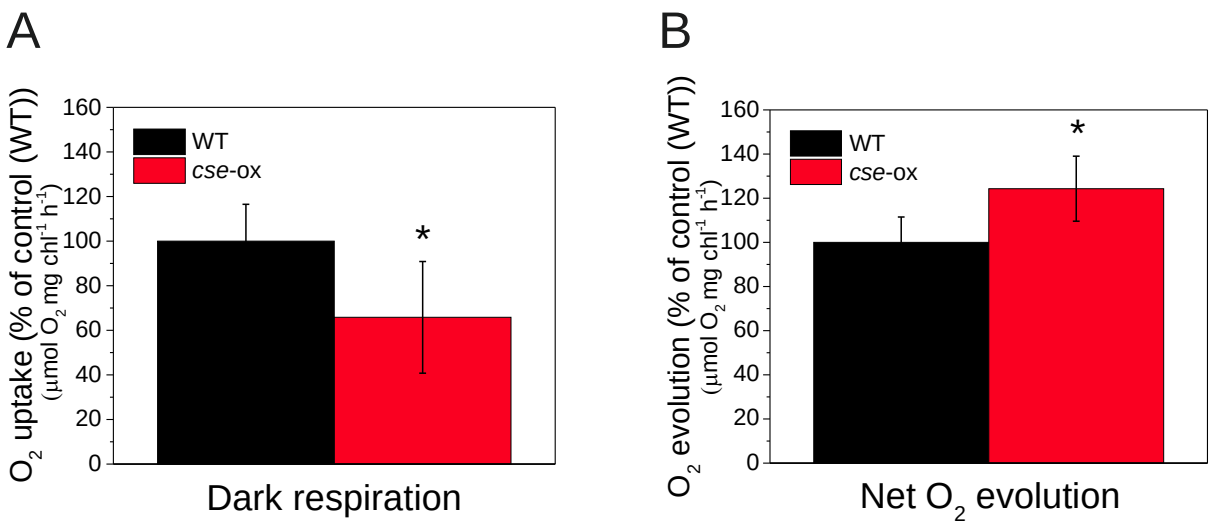


Figure 7 Oxygen exchange rates in 2 day-old *Anabaena* wild-type (WT) and *cse* over-expression (*cse*-ox) strains grown in $BG11_AC$

(A) Oxygen uptake in WT and *cse*-ox in darkness; **(B)** Net oxygen evolution in high light (400 µmol photons·m^{-2}·s^{-1}). Values represent means of four biological replicates, error bars show standard deviations. Significant differences between WT and mutant samples are indicated with asterisks (*t*-test $P<0.05$)

3.6 Abundance of protein complexes in the thylakoid membrane is affected by increased cse levels

To evaluate the composition of photosynthetic protein complexes, purified thylakoid membranes were solubilised with 1.5% DM, and protein complexes were separated by large pore blue native (lpBN)-PAGE using a 3.5 - 12.5% acrylamide gradient. Blue native PAGE separation revealed substantial increases in abundance of monomeric PSI and PSII complexes in the *cse*-ox strain compared to the WT, as well as an increase in monomeric PSII lacking the CP43 protein (Figure 8A). PSII dimers were virtually absent from cse-*ox*, and PSI dimers were also less abundant, while the PSI tetramer band appeared to be more prominent in *cse*-ox compared to WT. An unknown blue band migrating in the upper region of the lpBN-PAGE gel was substantially less abundant in *cse*-ox (see black arrow in Figure 8A). In the lower region of the BN gel, several blue bands appeared to have different migrations in *cse*-ox in comparison to WT (Figure 8A).

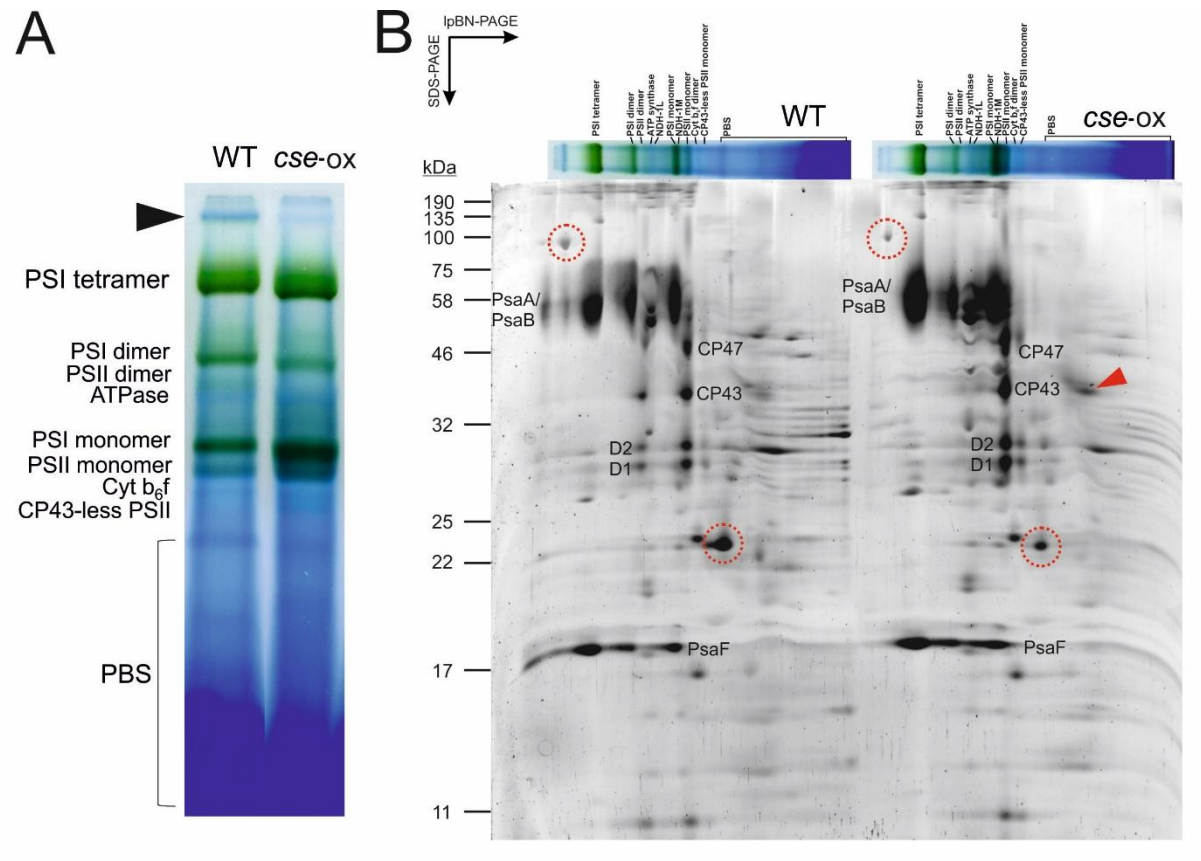


Figure 8 Photosynthetic thylakoid membrane protein complexes of 2 day-old *Anabaena* wild-type (WT) and *cse* over-expression (*cse*-ox) strains grown in $BG11_AC$

(A) Thylakoid membranes equivalent to 75 µg protein were solubilised with 1.5% *n*-dodecyl-β-D-maltoside (DM) and separated by large pore blue native-PAGE using an acrylamide gradient of 3.5 - 12.5%. Major protein complexes are labelled. Black arrow indicates unknown high molecular weight complex putatively associated with PSI (see text); **(B)** Separation of protein complexes in second dimension using SDS-PAGE containing 12% acrylamide. Proteins were stained with Sypro Ruby Protein Gel Stain. Characteristic subunits of PSI and PSII complexes are labelled. Red arrow indicates free CP43 protein. Protein spots cut from the WT gel are circled in red in both WT and *cse*-ox gels. Sizes of molecular weight markers are indicated on the left side

Subunit proteins of photosynthetic complexes separated by lpBN-PAGE were further separated according to their molecular mass using SDS-PAGE in the second dimension (Figure 8B). Protein spots corresponding to CP47, CP43, D1 and D2 subunits of PSII were clearly derived from both PSII dimers and monomers in WT samples, while these spots were virtually absent from positions corresponding to the larger PSII dimer complex in *cse*-ox and were more abundant in positions corresponding to the PSII monomer. PSII monomers lacking CP43 were also more abundant in *cse*-ox, as was a spot corresponding to free CP43 protein (see red arrow in Figure 8B). PsaA/B and PsaF subunits derived in denaturing PAGE from the upper region of the lpBN-PAGE gel (above the PSI tetramer band) revealed the existence of very high molecular weight PSI complexes in WT thylakoids that appeared to be absent from the *cse*-ox thylakoids. To investigate

whether these PSI complexes were associated with light-harvesting antennae, a 100 kDa protein spot in the SDS-PAGE gel (see Figure 8B), which corresponded to the high molecular weight band in the native gel described above, was cut from the WT gel and investigated by mass spectrometry (MS). A second protein spot of approximately 20 kDa, which was less abundant in *cse*-ox thylakoids, and appeared to migrate with a PBS band in the lpBN-PAGE, was also identified by MS (Figure 8B). "Unknown proteins" all3041 and alr2489 were identified with high confidence in the first and second protein spots, respectively. Both proteins were previously identified in *Anabaena* thylakoids [49], although no connection between these proteins and photosynthesis was apparent.

4. Discussion

Internal Ca^{2+} concentration ($[Ca^{2+}]_i$) in cyanobacteria is tightly controlled by ion channels/transporters and Ca^{2+}-binding proteins. Changes in free $[Ca^{2+}]_i$ regulate many cyanobacterial processes, including phototactic orientation, heterocyst differentiation and frequency, PBS degradation, hormogonia differentiation and gliding [reviewed in 50]. However, the mechanisms of Ca^{2+}-sensing and signalling in cyanobacteria are poorly understood. We describe here the discovery of CSE, a previously unreported EF-hand protein that binds Ca^{2+} *in vitro* and is highly conserved in filamentous, mostly heterocystous cyanobacteria. All CSE homologs identified in cyanobacteria genome databases contain an N-terminal EF-hand that perfectly matches the EF-hand consensus sequence [46]. The C-terminal EF-hand appears to have lower affinity for Ca^{2+} due to substitutions of key residues in this domain, which was supported by thermodynamic data that indicated a single Ca^{2+}-binding site (Table 1). Alterations in CSE structure upon Ca^{2+}-binding detected by CD analysis (Figure 4E) resembles the Ca^{2+}-mediated conformational changes of other EF-hand proteins such as calmodulin and troponin C [51], suggesting that Ca^{2+}-binding may regulate an interaction between CSE and receptor proteins as part of a Ca^{2+}-sensitive signalling pathway.

To investigate the function of CSE, an overexpression strain (*cse*-ox) was generated with a two-fold increase in *cse* expression compared to WT *Anabaena* (Table 3). This moderate increase in *cse* expression was sufficient to impair overall growth of the mutant strain, as well as causing yellow colouration of cultures and decreased levels of phycocyanobilin (PCB) pigment (Figure 3), which can indicate N deficiency and subsequent degradation of N-containing pigments [53]. A role for CSE in N metabolism was also alluded to by almost exclusive occurrence of *cse* gene homologues in filamentous, N-fixing cyanobacteria. Considering these findings together with previously established links between Ca^{2+} signalling and N step-down [30,53,54], we investigated whether the growth penalty in *cse*-ox may be due to impaired N metabolism. However, heterocyst frequency and nitrogenase activity in cells grown in media lacking combined N sources did not differ significantly from WT levels (Table 2). Instead, the observed decrease in growth was

consistent with a decrease in photosynthetic activity and substantial alterations in the composition of photosynthetic pigment-protein complexes, including the phycobilisome (PBS) light-harvesting antennae, and the multimeric states of PSII and PSI complexes. Notably, we also created knockout mutant *Anabaena* strains lacking the *cse* gene, which displayed many of the same phenotypical characteristics displayed by the *cse*-ox strain, including slow growth, altered photosynthetic activity and thylakoid membrane organisation. However, substantial variability was observed between different clones of the Δ*cse* knockout, in contrast to the phenotype of *cse*-ox that was consistent between several individual clones. Therefore, only the *cse*-ox strain is described here. Nonetheless, striking similarity between cells with moderate over-abundance of CSE and those lacking CSE supports the role of the CSE protein, and proper regulation of Ca^{2+}, as the underlying factor of the phenotypes described here, rather than possible artefactual causes related to the genetic modifications.

Transcriptomic analysis of *cse*-ox revealed strong downregulation of gene clusters encoding protein components of the light-harvesting complexes, in comparison to the WT. The entire *pec* gene cluster, which encodes the phycoerytherocyanin (PEC) proteins, and the *cpc* gene cluster, which encodes the phycocyanin (PC) proteins, were the most strongly downregulated genes in the mutant (Table 3). An exception was *cpcG4*, which is transcribed independently from the *cpc* cluster [20,27]. PC and PEC proteins make up the phycobiliprotein rods that radiate from the allophycocyanin (APC) core cylinders of the PBS complex. Unlike the peripheral PBS subunits, expression of genes encoding APC subunits was not affected in *cse*-ox, suggesting that PBS rods in *cse*-ox may be truncated. Disturbed PBS function in *cse*-ox was also apparent in fluorescence emission measurements at 77K and room temperature, both of which showed substantially lower emission from PC at 646 nm, and increased emission from APC at 663 nm, after PBS excitation with 580 nm light (Figure 4). Light energy absorbed by the PBS is normally transferred sequentially from PEC to PC, to APC, then to the terminal emitter ApcE and finally to P680 (PSII reaction centre chlorophyll) to initiate linear electron transport (LET) towards PSI [22]. Reduced PC fluorescence emission was consistent with lower PC and PEC abundance, while enhanced fluorescence emission from APC (Figure 4) indicates inefficient energy transfer between PBS and the photosystem reaction centres, resulting in release of excitation directly from APC. Increased APC fluorescence has previously been linked to abnormal PBS rod composition [24].

PSII:PSI ratios were equivalent between WT and the *cse*-ox mutant (Figure 4A), indicating that defective excitation transfer from PBS was not due to any changes in photosystem stoichiometry in the mutant. On the other hand, the virtual absence of PSII complex dimers (Figure 8A), which are required for PSII-PBS connectivity [24,55-59], most likely had a substantial negative impact on normal excitation transfer from PBS. Corresponding increases in PSII monomers, both with and without the CP43 subunit, as well as free CP43 protein, were clearly evident in *cse*-ox thylakoids

(Figure 8). This finding indicates a high rate of PSII damage and turn-over in the mutant [60,61], which may be caused by over-reduction of the photosynthetic electron carriers and increased excitation pressure on PSII. Indeed, such over-reduction was evident in high levels of chlorophyll *a* fluorescence indicative of more closed PSII centres in both darkness and actinic light, and in the slower rate of fluorescence decay phase that is attributed to slower binding of oxidised PQ to the Q_B pocket (see Figure 5D; 0.006 – 3 s). A steady decrease in PSII efficiency observed in *cse*-ox exposed to "state 1 light" (Figure 5F) may have been caused by progressive PSII damage under light conditions that induce PBS association with PSII. Oddly, PSII damage may have been exacerbated in the mutant by downregulated expression of genes encoding three helical carotenoid proteins (HCPs) that are homologous to the OCP N-terminal domain [62]. Two of the downregulated HCPs are active in quenching of singlet oxygen ($^1O_2$) [63], suggesting a decreased $^1O_2$ scavenging capacity that may have enhanced the level of PSII damage in *cse*-ox. The conventional OCP gene (*all3149*) that quenches $^1O_2$ as well as excess excitation directly from PBS under light stress [64,65], was not found to be differentially expressed in our data.

Altered PSI activity was detected in *cse*-ox, including re-reduction of oxidised $P700^+$ during far-red (FR) light illumination, lower Y(ND) after FR exposure, and higher Y(NA) under actinic light (Figure 5). These effects can also be attributed to over-reduction of photosynthetic electron carriers in the mutant leading to an abnormally high flow of electrons to the PSI donor side. Clearly this over-reduction could not be alleviated by PSI activity, in spite of an apparent increase in PSI complex abundance evident in *cse*-ox thylakoids (Figure 8). It is possible that inefficient excitation transfer to PSI from abnormal PBS lacking PC and PEC (discussed above) may have inhibited PSI electron transport (Figure 8).

Inhibited respiration in *cse*-ox was evident in the lower rate of oxygen uptake in darkness (Figure 7A), and in the higher rate of net oxygen evolution in the light (Figure 7B). Considering the lower PSII activity in *cse*-ox (described above), higher net oxygen evolution is unlikely to have been caused by increased PSII water-splitting, instead implying lower oxygen consumption in the over-expressor in the light. Respiration in cyanobacteria provides several alternative pathways of electron transport from metabolites, via PQ, to terminal oxidases in the thylakoid and plasma membranes, operating in both light and darkness [reviewed in 66]. Decreased electron flux through respiratory pathways could explain over-reduction of photosynthetic electron carriers and the subsequent negative effects on PSII and PSI activity described above. Indeed, higher chlorophyll fluorescence and over-reduced P700 were observed in mutant cyanobacteria lacking the respiratory terminal oxidase (RTO) enzymes cytochrome *bd* quinol oxidase (Cyd) and cytochrome *c* oxidase (COX) [67]. The results observed in *cse*-ox could also be attributed to decreased oxygen photoreduction by the flavodiiron proteins, which are especially important for protection of PSII and PSI in stress conditions [68].

5. Conclusions and perspectives

The results presented here show the newly discovered CSE protein to be a *bona fide* Ca^{2+}-binding EF-hand protein. Therefore, disrupted electron transport and thylakoid protein complex formation/stability in *Anabaena* with over-abundance of CSE is thought to relate to modified Ca^{2+} signalling and/or a change in Ca^{2+} homeostasis. Although the precise mechanism behind the observed phenotypes of *cse*-ox has not yet been identified, several factors that regulate photosynthetic energy balance in cyanobacteria have established links with Ca^{2+}. The flavodiiron protein Flv4 contains Ca^{2+}-binding sites, and the presence of cations (Ca^{2+} or Mg^{2+}) was shown to influence partitioning of the Flv2-Flv4 dimer between the thylakoid membrane (cations present) and the soluble fraction (cations absent) [69]. A mutant strain lacking *flv4* showed substantial phenotypic overlap with the *cse*-ox mutant, including increased PSII monomer:dimer ratio, over-reduction of the PQ pool and higher fluorescence emission from detached APC [69-71]. Furthermore, expression of *flv4* in *Anabaena* was downregulated by a shift to low Ca^{2+} conditions [30]. It is possible that over-expression of *cse* may interfere with Flv4 association with the thylakoid, leading to PQ over-reduction and increased PSII monomerisation. A Ca^{2+}-binding site is conserved in prokaryotic and eukaryotic orthologues of subunit A of the aa3-type COX terminal oxidase [72-74]. This site appears to be also present in the predicted *Anabaena* COXA structure (see supplemental Figure S4). COX activity was downregulated by Ca^{2+}-binding to mitochondrial COXA [75], suggesting that abnormally high or otherwise unregulated Ca^{2+} levels in *cse*-ox may have led to decreased respiration activity and subsequent over-reduction of PQ through inhibition of COX [67].

The Ca^{2+}-induced conformational change in CSE, taken together with the activity of other EF-hand proteins such as calmodulin and troponin C, suggests that CSE is likely to interact with partner proteins in a Ca^{2+}-dependent manner. Interaction partners may include protein phosphatases, such as PrpA which was upregulated in *cse*-ox, or protein kinases. PBS proteins are known to be phosphorylated when assembled [76,77], and so dysfunctional Ca^{2+}-dependent regulation of PBS phosphorylation may explain the modified PBS composition identified here. Furthermore, abnormally-abundant proteins in *cse*-ox identified by MS (all3041 and alr2489) were found to be heavily phosphorylated in the WT (data not shown). Identification of interaction partners will be an important step in understanding the molecular mechanism of the CSE protein.

An outstanding question relates to the evolution of CSE, which appears almost exclusively in filamentous, N-fixing cyanobacteria. Changes in $[Ca^{2+}]_i$ in *Anabaena* have been linked to the differentiation of vegetative cells into heterocysts under N-deficient conditions [48,53], and appear to play a role in communicating C:N homeostasis [30]. The current work shows that CSE over-abundance does not affect normal heterocyst differentiation and nitrogenase activity; however, the

observed effects on photosynthetic activity may be expected to downregulate carbon metabolism and thereby impact C:N balance. During heterocyst differentiation, substantial changes to the organisation and function of thylakoid membranes are known to occur, including decreased abundance of PBS [78] and increased respiration via COX [79,80], which may link the photosynthetic phenotypes reported here with disrupted Ca^{2+} signalling in filamentous strains.

6. Acknowledgements

J.W. and P.J.G thank Ilaria Mancini for running lpBN-PAGE gels, Andrea Trotta and Azfar Ali Bajwa for protein MS analysis, Yagut Allahverdiyeva, Natalia Battchikova and Mikko Tikkanen for help with data analysis and useful discussions. K.A.S. thanks Marcus Hartmann and Reinhard Albrecht for the assistance with CD spectroscopy. Financial support was provided from EU Marie Curie ITN CALIPSO project GA ITN 2013–607607 (J.W.) and Academy of Finland projects 307335 and 303757 (E-M.A.) and 26080341 (P.J.G.). K.A.S and K.F. were supported by Deutsche Forschungsgemeinschaft (DFG).

6. References

6. References

1. Clapham, D.E. (2007) Calcium signaling. *Cell* **131**(6): 1047-1058. doi: 10.1016/j.cell.2007.11.028
2. Norris, V., Grant, S., Freestone, P., Canvin, J., Sheikh, F.N., Toth, I., Trinei, M., *et al.* (1996) Calcium signalling in bacteria. *J Bacteriol* **178**(13): 3677-3682. doi: 10.1128/jb.178.13.3677-3682.1996
3. Domínguez, D.C., Guragain, M., and Patrauchan, M. (2015) Calcium binding proteins and calcium signaling in prokaryotes. *Cell Calcium* **57**(3): 151-165. doi: 10.1016/j.ceca.2014.12.006
4. Stael, S., Wurzinger, B., Mair, A., Mehlmer, N., Vothknecht, U.C., and Teige, M. (2012) Plant organellar calcium signaling: an emerging field. *J Exp Bot* **63**(4): 1525-1542. doi: 10.1093/jxb/err394
5. Nomura, H., and Shiina, T. (2014) Calcium signaling in plant endosymbiotic organelles: mechanism and role in physiology. *Mol Plant* **7**(7): 1094-1104. doi: 10.1093/mp/ssu020
6. Hochmal, A.K., Schulze, S., Trompelt, K., and Hippler, M. (2015) Calcium-dependent regulation of photosynthesis. *Biochim Biophys Acta* **1847**(9): 993-1003. doi: 10.1016/j.bbabio.2015.02.010
7. Kmiecik, P., Leonardelli, M., and Teige, M. (2016) Novel connections in plant organellar signalling link different stress responses and signalling pathways. *J Exp Bot* **67**(13): 3793-3807. doi: 10.1093/jxb/erw136

8. Gould, S.B., Waller, R.F., and McFadden, G.I. (2008) Plastid evolution. *Annu Rev Plant Biol* **59**: 491-517. doi: 10.1146/annurev.arplant.59.032607.092915
9. Giraldez-Ruiz, N., Mateo, P., Bonilla, I., and Fernández-Piñas, F. (1997) The relationship between intracellular pH, growth characteristics and calcium in the cyanobacterium *Anabaena* sp. strain PCC7120 exposed to low pH. *New Phytol* **137**: 599-605. Retrieved from http://www.jstor.org/stable/2558969
10. Torrecilla, I., Leganés, F., Bonilla, I., and Fernández-Piñas, F. (2004) Light-to-dark transitions trigger a transient increase in intracellular Ca^{2+} modulated by the redox state of the photosynthetic electron transport chain in the cyanobacterium *Anabaena* sp. PCC7120. *Plant Cell Environ* **27**(7): 810-819. doi: 10.1111/j.1365-3040.2004.01187.x
11. Sai, J., and Johnson, C.H. (2002) Dark-stimulated calcium ion fluxes in the chloroplast stroma and cytosol. *Plant Cell* **14**(6): 1279-1291. doi: 10.1105/tpc.000653
12. Fredricks, W.W., and Jagendorf, A.T. (1964) A soluble component of the Hill reaction in *Anacystis nidulans*. *Arch Biochem Biophys* **104**: 39-49. doi: 10.1016/S0003-9861(64)80032-1
13. Piccioni, R.G., and Mauzerall, D.C. (1978) Calcium and photosynthetic oxygen evolution in cyanobacteria. *Biochim Biophys Acta* **504**(3): 384-397. doi: 10.1016/0005-2728(78)90061-0
14. Piccioni, R.G., and Mauzerall, D.C. (1976) Increase effected by calcium ion in the rate of oxygen evolution from preparations of *Phormidium luridum*. *Biochim Biophys Acta* **423**(3): 605-609. doi: 10.1016/0005-2728(76)90213-9
15. Brand, J.J., Mohanty, P., and Fork, D.C. (1983) Reversible inhibition of photochemistry of photosystem II by Ca2+ removal from intact cells of *Anacystis nidulans*. *FEBS Lett*. **155**(1): 120-124. doi: 10.1016/0014-5793(83)80222-1
16. Satoh, K., and Katoh, S. (1985) A functional site of Ca^{2+} in the oxygen-evolving photosystem II preparation from *Synechococcus* sp. *FEBS* **190**(2): 199-203. doi: 10.1016/0014-5793(85)81283-7
17. Müh, F., Renger, T., and Zouni, A. (2008) Crystal structure of cyanobacterial photosystem II at 3.0 A resolution: a closer look at the antenna system and the small membrane-intrinsic subunits. *Plant Physiol Biochem* **46**(3): 238-264. doi: 10.1016/j.plaphy.2008.01.003
18. Umena, Y., Kawakami, K., Shen, J.R., and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**(7345): 55-60. doi: 10.1038/nature09913
19. Malavath, T., Caspy, I., Netzer-El, S.Y., Klaiman, D., and Nelson, N. (2018) Structure and function of wild-type and subunit-depleted photosystem I in *Synechocystis*. *Biochim Biophys Acta Bioenerg* **1859**(9): 645-654. doi: 10.1016/j.bbabio.2018.02.002
20. Watanabe, M., Semchonok, DA., Webber-Birungi, M.T., Ehira, S., Kondo, K., Narikawa, R., Ohmori, M., Boekema, E.J., Ikeuchi, M. (2014) Attachment of phycobilisomes in an

antenna-photosystem I supercomplex of cyanobacteria. *Proc Natl Acad Sci USA* **111**(7): 2512-2517. doi: 10.1073/pnas.1320599111

21. Liu, H., Zhang, H., Niedzwiedzki, D.M., Prado, M., He, G., Gross, M.L., and Blankenship, R.E. (2013) Phycobilisomes supply excitations to both photosystems in a megacomplex in cyanobacteria. *Science* **342**(6162): 1104-1107. doi: 10.1126/science.1242321
22. Stadnichuk, I.N., Krasilnikov, P.M., and Zlenko, D.V. (2015) Cyanobacterial phycobilisomes and phycobiliproteins. *Microbiology* **84**(2): 101-111. doi: 10.1134/S0026261715020150
23. Bryant, D.A. (1982) Phycoerythrocyanin and phycoerythrin: Properties and occurrence in cyanobacteria. *J Gen Microbiol* **128**: 835-844. doi: 10.1099/00221287-128-4-835
24. Chang, L., Liu, X., Li, Y., Liu, C.C., Yang, F., Zhao, J., and Sui, S.F. (2015) Structural organization of an intact phycobilisome and its association with photosystem II. *Cell Res* **25**(6): 726-737. doi: 10.1038/cr.2015.59
25. Tang, K., Ding, W.L., Höppner, A., Zhao, C., Zhang, L., Hontani, Y., Kennis, J.T., Gärtner, W., Scheer, H., Zhou, M., Zhao, K.H. (2015) The terminal phycobilisome emitter, LCM: A light-harvesting pigment with a phytochrome chromophore. *Proc Natl Acad Sci U S A* **112**(52): 15880-15885. doi: 10.1073/pnas.1519177113
26. Dong, C., Tang, A., Zhao, J., Mullineaux, C.W., Shen, G., and Bryant, D.A. (2009) ApcD is necessary for efficient energy transfer from phycobilisomes to photosystem I and helps to prevent photoinhibition in the cyanobacterium *Synechococcus* sp. PCC 7002. *Biochim Biophys Acta* **1787**(9): 1122-1128. doi: 10.1016/j.bbabio.2009.04.007
27. Bryant, D.A., Stirewalt, V.L., Glauser, M., Frank, G., Sidler, W., and Zuber, H. (1991) A small multigene family encodes the rod-core linker polypeptides of *Anabaena* sp. PCC7120 phycobilisomes. *Gene* **107**(1): 91-99. doi: 10.1016/0378-1119(91)90301-Q
28. Ducret, A., Müller, S.A., Goldie, K.N., Hefti, A., Sidler, W.A., Zuber, H., and Engel, A. (1998) Reconstitution, characterisation and mass analysis of the pentacylindrical allophycocyanin core complex from the cyanobacterium *Anabaena* sp. PCC 7120. *J Mol Biol* **278**(2): 369-388. doi: 10.1006/jmbi.1998.1678
29. Muro-Pastor, A.M., Herrero, A. and Flores, E. (2001) Nitrogen-Regulated Group 2 Sigma Factor fromSynechocystis sp. Strain PCC 6803 Involved in Survival under Nitrogen Stress. *Journal of bacteriology*, **183**(3), pp.1090-1095.
30. Walter, J., Lynch, F., Battchikova, N., Aro, E.M., and Gollan, P.J. (2016) Calcium impacts carbon and nitrogen balance in the filamentous cyanobacterium *Anabaena* sp. PCC 7120. *J Exp Bot* **67**(13): 3997-4008. doi: 10.1093/jxb/erw112
31. Wolk, C.P., Fan, Q., Zhou, R., Huang, G., Lechno-Yossef, S., Kuritz, T., and Wojciuch, E. (2007) Paired cloning vectors for complementation of mutations in the cyanobacterium Anabaena sp. strain PCC 7120. Arch Microbiol 188(6): 551-563. doi: 10.1007/s00203-007-0276-z

32. Elhai, J., and Wolk, C.P. (1988b) Conjugal transfer of DNA to cyanobacteria. *Methods Enzymol* **167**: 747-754. doi: 10.1016/0076-6879(88)67086-8
33. Stael, S., Rocha, A.G., Robinson, A.J., Kmiecik, P., Vothknecht, U.C., and Teige, M. (2011) *Arabidopsis* calcium-binding mitochondrial carrier proteins as potential facilitators of mitochondrial ATP-import and plastid SAM-import. *FEBS Lett* **585**(24): 3935-3940. doi: 10.1016/j.febslet.2011.10.039
34. Lapina, T., Selim, K.A., Forchhammer, K., and Ermilova, E. (2018) The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins. *Sci Rep* **8**: 790. doi: 10.1038/s41598-017-19046-7
35. Selim, K.A., Haase, F., Hartmann, M.D., Hagemann, M., and Forchhammer, K. (2018) PII-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. *Proc Natl Acad Sci U S A* **115**(21): E4861-E4869. doi: 10.1073/pnas.1803790115
36. Hauf, K., Kayumov, A., Gloge, F., and Forchhammer, K. (2016) The molecular basis of TnrA control by glutamine synthetase in *Bacillus subtilis*. *J Biol Chem* **291**(7): 3483-3495. doi: 10.1074/jbc.M115.680991
37. Korycinski, M., Albrecht, R., Ursinus, A., Hartmann, M.D., Coles, M., Martin, J., Dunin-Horkawicz, S., Lupas, A.N. (2015) STAC - A new domain associated with transmembrane solute transport and two-component signal transduction systems. *J Mol Biol* **427**(20): 3327-3339. doi: 10.1016/j.jmb.2015.08.017
38. Kosourov, S., Leino, H., Murukesan, G., Lynch, F., Sivonen, K., Tsygankov, A.A., Aro, E.M., *et al.* (2014) Hydrogen photoproduction by immobilized n2-fixing cyanobacteria: understanding the role of the uptake hydrogenase in the long-term process. *Appl Environ Microbiol* **80**(18): 5807-5817. doi: 10.1128/AEM.01776-14
39. Kämäräinen, J., Huokko, T., Kreula, S., Jones, P.R., Aro, E.M., and Kallio, P. (2016) Pyridine nucleotide transhydrogenase PntAB is essential for optimal growth and photosynthetic integrity under low-light mixotrophic conditions in *Synechocystis* sp. PCC 6803. *New Phytol* **214**(1): 194-204. doi: 10.1111/nph.14353
40. Huokko, T., Muth-Pawlak, D., Battchikova, N., Allahverdiyeva, Y., and Aro, E.M. (2017) Role of type 2 NAD(P)H dehydrogenase NdbC in redox regulation of carbon allocation in *Synechocystis*. *Plant Physiol* **174**(3): 1863-1880. doi: 10.1104/pp.17.00398
41. Zhang, P., Allahverdiyeva, Y., Eisenhut, M., and Aro, E.M. (2009) Flavodiiron proteins in oxygenic photosynthetic organisms: photoprotection of photosystem II by Flv2 and Flv4 in *Synechocystis* sp. PCC 6803. *PLoS One* **4**(4): e5331. doi: 10.1371/journal.pone.0005331
42. Trotta, A., Suorsa, M., Rantala, M., Lundin, B. and Aro, E.M. (2016) Serine and threonine residues of plant STN 7 kinase are differentially phosphorylated upon changing light

conditions and specifically influence the activity and stability of the kinase. *The Plant Journal*, **87**(5), pp.484-494.

43. Fewer, D., Friedl, T., and Büdel, B. (2002) *Chroococcidiopsis* and heterocyst-differentiating cyanobacteria are each other's closest living relatives. *Mol Phylogenet Evol* **23**(1): 82-90. doi: 10.1006/mpev.2001.1075
44. Rigonato, J., Gama W.A., Alvarenga, D.O., Branco, L.H., Brandini, F.P., Genuário, D.B., and Fiore, M.F. (2016) *Aliterella atlantica* gen. nov., sp. nov., and *Aliterella antarctica* sp. nov., novel members of coccoid cyanobacteria. *Int J Syst Evol Microbiol* **66**(8): 2853-2861. doi: 10.1099/ijsem.0.001066
    0373-3
45. Bergman, B., Gallon, J.R., Rai, A.N., and Stal, L.J. (1997) N2 Fixation by non-heterocystous cyanobacteria. *FEMS Microbiol Rev* **19**: 139-185. doi: 10.1111/j.1574-6976.1997.tb00296.x
46. Michiels, J., Xi, C., Verhaert, J., and Vanderleyden, J. (2002) The functions of Ca(2+) in bacteria: a role for EF-hand proteins? *Trends Microbiol* **10**(2): 87-93. doi: 10.1016/S0966-842X(01)02284-3
47. Kretsinger, R.H., and Nockolds, C.E. (1973) Carp muscle calcium-binding protein. II. Structure determination and general description. *J Biol Chem* **248**(9): 3313-3326. Retrieved from http://www.jbc.org/content/248/9/3313
48. Hu, Y., Zhang, X., Shi, Y., Zhou, Y., Zhang, W., Su, X.D., Xia, B., *et al.* (2011) Structures of *Anabaena* calcium-binding protein CcbP: insights into Ca2+ signaling during heterocyst differentiation. *J Biol Chem* **286**(14): 12381-12388. doi: 10.1074/jbc.M110.201186
49. Ladig, R., Sommer, M.S., Hahn, A., Leisegang, M.S., Papasotiriou, D.G., Ibrahim, M., Elkehal, R., Karas, M., Zickermann, V., Gutensohn, M. and Brandt, U. (2011) A high-definition native polyacrylamide gel electrophoresis system for the analysis of membrane complexes. *The Plant Journal*, **67**(1), pp.181-194.
50. Agostoni, M., and Montgomery, B.L. (2014) Survival strategies in the aquatic and terrestrial world: the impact of second messengers on cyanobacterial processes. *Life* **4**(4): 745-769. doi: 10.3390/life4040745ä
51. Denessiouk, K., Permyakov, S., Denesyuk, A., Permyakov, E., and Johnson, M.S. (2014) Two structural motifs within canonical EF-hand calcium-binding domains identify five different classes of calcium buffers and sensors. *PLoS One* **9**(10): e109287. doi: 10.1371/journal.pone.0109287
52. Schwarz, R., and Forchhammer, K. (2005) Acclimation of unicellular cyanobacteria to macronutrient deficiency: emergence of a complex network of cellular responses. *Microbiology* **151**(8): 2503-2514. doi: 10.1099/mic.0.27883-0

53. Torrecilla, I., Leganés, F., Bonilla, I., and Fernández-Piñas, F. (2004) A calcium signal is involved in heterocyst differentiation in the cyanobacterium *Anabaena* sp. PCC7120. *Microbiology* **150**(11): 3731-3739. doi: 10.1099/mic.0.27403-0
54. Leganés, F., Forchhammer, K., and Fernández-Piñas, F. (2009) Role of calcium in acclimation of the cyanobacterium *Synechococcus elongatus* PCC 7942 to nitrogen starvation. *Microbiology* **155**(Pt 1): 25-34. doi: 10.1099/mic.0.022251-0
55. Mörschel, E., and Mühlethaler, K. (1983) On the linkage of exoplasmatic freeze-fracture particles to phycobilisomes. *Planta* **158**(5): 451-457. doi: 10.1007/BF00397739
56. Mörschel, E., and Schatz, G.H. (1987) Correlation of photosystem-II complexes with exoplasmatic freeze-fracture particles of thylakoids of the cyanobacterium *Synechococcus* sp. *Planta* **172**(2): 145-154. doi: 10.1007/BF00394582
57. Bald, D., Kruip, J., and Rögner, M. (1996) Supramolecular architecture of cyanobacterial thylakoid membranes: How is the phycobilisome connected with the photosystems? *Photosynth Res* **49**(2): 103-118. doi: 10.1007/BF00117661
58. Barber, J., Morris, E.P., and da Fonseca, P.C. (2003) Interaction of the allophycocyanin core complex with photosystem II. *Photochem Photobiol Sci* **2**(5): 536-541. doi: 10.1039/B300063J
59. Arteni, A.A., Ajlani, G., and Boekema, E.J. (2009) Structural organisation of phycobilisomes from *Synechocystis* sp. strain PCC6803 and their interaction with the membrane. *Biochim Biophys Acta* **1787**(4): 272-279. doi: 10.1016/j.bbabio.2009.01.009
60. Aro, E.M., Suorsa, M., Rokka, A., Allahverdiyeva, Y., Paakkarinen, V., Saleem, A., Battchikova, N., Rintamäki, E. (2005) Dynamics of photosystem II: a proteomic approach to thylakoid protein complexes. *J Exp Bot* **56**(411): 347-356. doi: 10.1093/jxb/eri041
61. Nixon, P.J., Michoux, F., Yu, J., Boehm, M., and Komenda, J. (2010) Recent advances in understanding the assembly and repair of photosystem II. *Ann Bot* **106**(1): 1-16. doi: 10.1093/aob/mcq059
62. Melnicki, M.R., Leverenz, R.L., Sutter, M., López-Igual, R., Wilson, A., Pawlowski, E.G., Perreau, F., Kirilovsky, D., Kerfeld, C.A. (2016) Structure, diversity, and evolution of a new family of soluble carotenoid-binding proteins in cyanobacteria. *Molecular plant*, **9**(10), pp.1379-1394.
63. López-Igual, R., Wilson, A., Leverenz, R.L., Melnicki, M.R., Bourcier de Carbon, C., Sutter, M., Turmo, A., Perreau, F., Kerfeld, C.A., Kirilovsky, D. (2016) Different functions of the paralogs to the N-terminal domain of the Orange Carotenoid Protein in the cyanobacterium *Anabaena* sp. PCC 7120. *Plant Physiol* **171**(3): 1852-1866. doi: 10.1104/pp.16.00502
64. Wilson, A., Ajlani, G., Verbavatz, J.M., Vass, I., Kerfeld, C.A., and Kirilovsky, D. (2006) A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria. *Plant Cell* **18**(4): 992-1007. doi: 10.1105/tpc.105.040121

65. Sedoud, A., López-Igual, R., ur Rehman, A., Wilson, A., Perreau, F., Boulay, C., Vass, I., Krieger-Liszkay, A., Kirilovsky, D. (2014) The cyanobacterial photoactive orange carotenoid protein is an excellent singlet oxygen quencher. *The Plant Cell*, pp.tpc-114.
66. Lea-Smith, D.J., Bombelli, P., Vasudevan, R., Howe, C.J., 2016. Photosynthetic, respiratory and extracellular electron transport pathways in cyanobacteria. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1857(3), pp.247-255.
67. Ermakova, M., Huokko, T., Richaud, P., Bersanini, L., Howe, C.J., Lea-Smith, D., Peltier, G., Allahverdiyeva, Y. (2016) Distinguishing the roles of thylakoid respiratory terminal oxidases in the cyanobacterium Synechocystis sp. PCC 6803. Plant physiology **171** 1307-1319. doi: https://doi.org/10.1104/pp.16.00479
68. Allahverdiyeva, Y., Isojärvi, J., Zhang, P. and Aro, E.M. (2015) Cyanobacterial oxygenic photosynthesis is protected by flavodiiron proteins. Life, 5(1), pp.716-743.
69. Zhang, P., Eisenhut, M., Brandt, A.M., Carmel, D., Silén, H.M., Vass, I., Allahverdiyeva, Y., *et al.* (2012) Operon *flv4-flv2* provides cyanobacterial photosystem II with flexibility of electron transfer. *Plant Cell* **24**(5): 1952-1971. doi: 10.1105/tpc.111.094417
70. Bersanini, L., Battchikova, N., Jokel, M., Rehman, A., Vass, I., Allahverdiyeva, Y., and Aro, E.M. (2014) Flavodiiron protein Flv2/Flv4-related photoprotective mechanism dissipates excitation pressure of PSII in cooperation with phycobilisomes in cyanobacteria. *Plant Physiol* **164**(2): 805-818. doi: 10.1104/pp.113.231969
71. Chukhutsina, V., Bersanini, L., Aro, E.M., and van Amerongen, H. (2015) Cyanobacterial flv4-2 operon-encoded proteins optimize light harvesting and charge separation in Photosystem II. *Mol Plant* **8**(5): 747-761. doi: 10.1016/j.molp.2014.12.016
72. Wikström, M.K. and Saari, H.T. (1977) The mechanism of energy conservation and transduction by mitochondrial cytochrome c oxidase. *Biochimica et Biophysica Acta* (BBA)-Bioenergetics, **462**(2), pp.347-361.
73. Yoshikawa, S., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., Yamashita, E., Inoue, N., Yao, M., Fei, M.J., Libeu, C.P., Mizushima, T., Yamaguchi, H. (1998) Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase. *Science*, 280(5370), pp.1723-1729.
74. Svensson-Ek, M., Abramson, J., Larsson, G., Törnroth, S., Brzezinski, P., Iwata, S., 2002. The X-ray crystal structures of wild-type and EQ (I-286) mutant cytochrome c oxidases from Rhodobacter sphaeroides. *Journal of molecular biology*, **321**(2), pp.329-339.
75. Vygodina, T., Kirichenko, A. and Konstantinov, A.A. (2013) Direct regulation of cytochrome c oxidase by calcium ions. PloS one, 8(9), p.e74436.
76. Piven, I., Ajlani, G., and Sokolenko, A. (2005) Phycobilisome linker proteins are phosphorylated in *Synechocystis* sp. PCC 6803. *J Biol Chem* **280**(22): 21667-21672. doi: 10.1074/jbc.M412967200

77. Angeleri, M., Muth-Pawlak, D., Aro, E.M., and Battchikova, N. (2016) Study of O-phosphorylation sites in proteins involved in photosynthesis-related processes in *Synechocystis* sp. strain PCC 6803: Application of the SRM approach. *J Proteome Res* **15**(12): 4638-4652. doi: 10.1021/acs.jproteome.6b00732
78. Wolk, C.P. and Simon, R.D. (1969) Pigments and lipids of heterocysts. Planta, **86**(1), pp.92-97.
79. Valladares, A., Maldener, I., Muro-Pastor, A.M., Flores, E. and Herrero, A., 2007. Heterocyst development and diazotrophic metabolism in terminal respiratory oxidase mutants of the cyanobacterium Anabaena sp. strain PCC 7120. Journal of bacteriology, 189(12), pp.4425-4430.
80. Torrado, A., Ramírez-Moncayo, C., Navarro, J.A., Mariscal, V. and Molina-Heredia, F.P. (2019) Cytochrome c6 is the main respiratory and photosynthetic soluble electron donor in heterocysts of the cyanobacterium Anabaena sp. PCC 7120. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, **1860**(1): 60-68.

Supplementary figures and tables

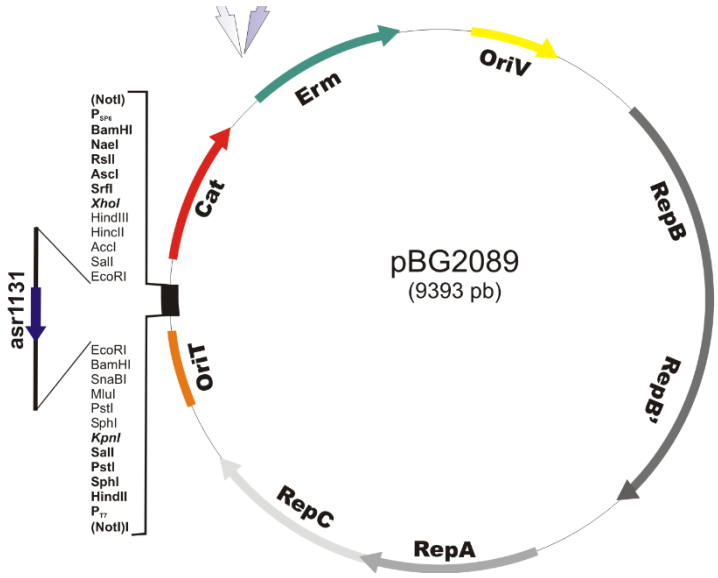


Figure S1: *cse* overexpression vector *pBG2089*

The *asr1131* gene under the control of its native promoter was cloned into the plasmid *pRL1342*, a *RSF1010* derivative designed for very low copy numbers, generating plasmid *pBG2089*.

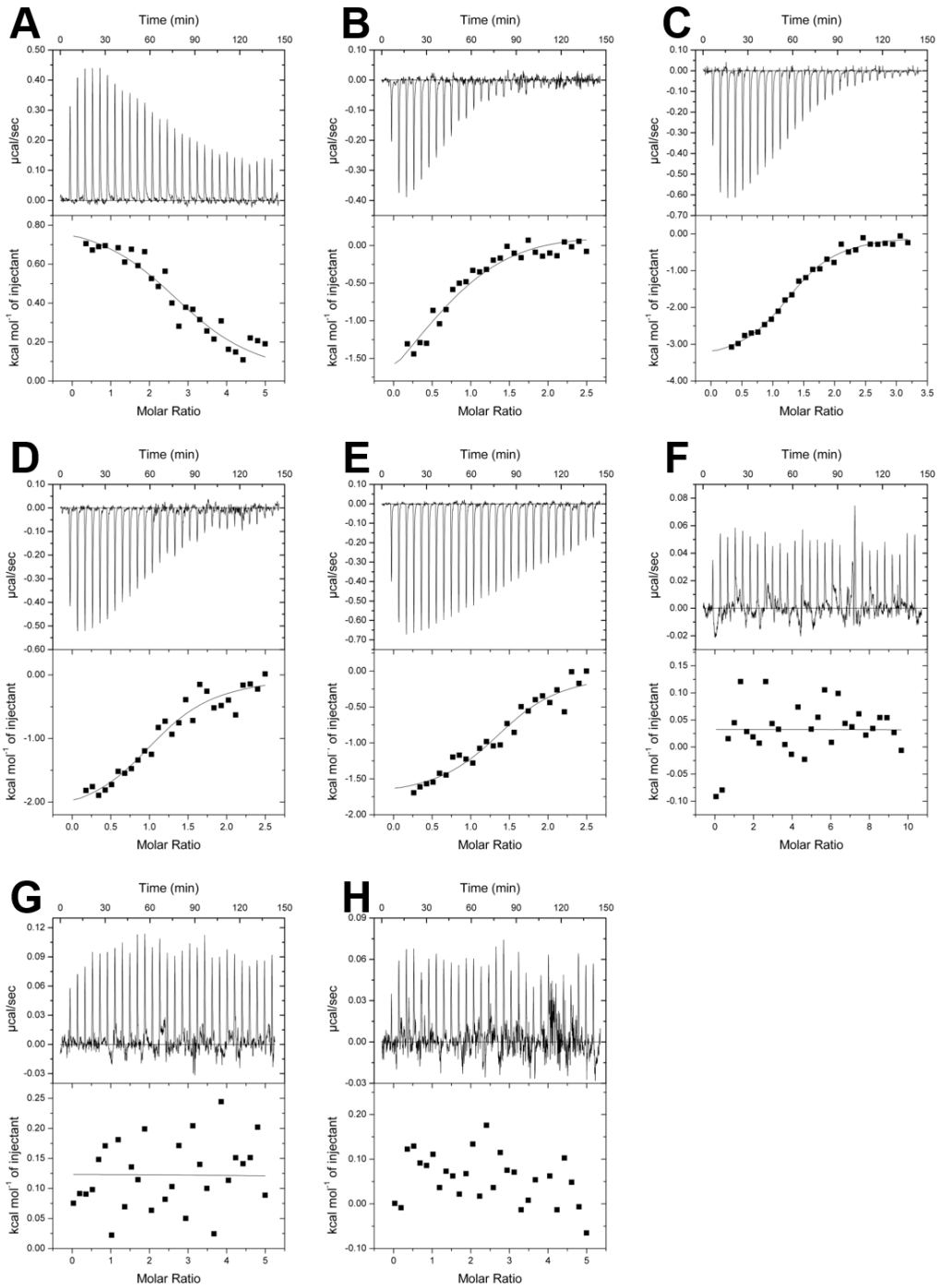


Figure S2: Isothermal titration calorimetry (ITC) metal-binding assay of recombinant His-tagged CSE in different buffers.

A. 85 µM CSE in 20 mM Tris buffer against 2 mM $MgCl_2$ **B.** 85 µM CSE in 20 mM Tris buffer + 100 mM NaCl against 1 mM $CaCl_2$ **C.** 66 µM CSE in 20 mM Tris buffer + 300 µM $MgCl_2$ against 0.75 mM $CaCl_2$ **D.** 85 µM CSE in 20 mM Tris buffer + 300 µM $MgCl_2$ + 100 mM NaCl against 1 mM $CaCl_2$ **E.** 85 µM CSE in 20 mM Tris buffer + 2 mM $MgCl_2$ + 100 mM NaCl against 1 mM $CaCl_2$ **F.** 66 µM CSE in 20 mM Tris buffer + 150 µM $CaCl_2$ against 3 mM $MgCl_2$ **G.** 85 µM CSE in 20 mM Tris buffer + 100 mM NaCl against 2 mM $MgCl_2$ **H.** 85 µM CSE in 20 mM Tris buffer + 100 mM NaCl + 150 µM $CaCl_2$ against 2 mM $MgCl_2$

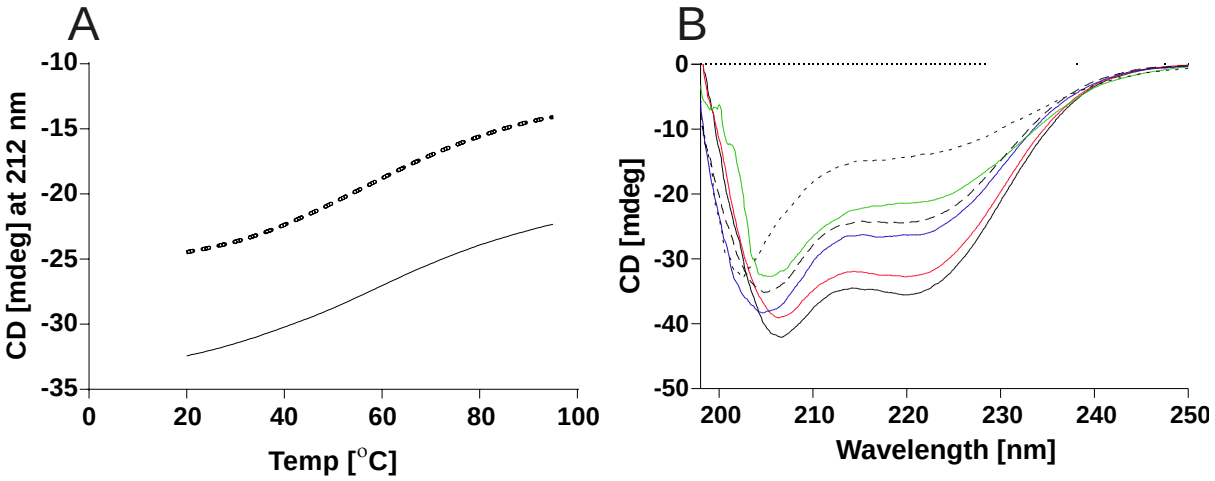


Figure S3: Circular dichroism (CD) spectra of recombinant His-tagged CSE in different buffers.

A. CSE melting curves from 20-95°C at 212 nm in 20 mM Tris buffer (dashed line) and 20 mM Tris buffer + 100 mM NaCl + 1 mM $CaCl_2$ (solid line) **B.** CD spectra of CSE in 20 mM Tris buffer at 20°C (after melting; blue solid line), at 20°C (black dashed line), at 95°C (black dotted line), and in 20 mM Tris buffer + 100 mM NaCl + 1 mM $CaCl_2$ at 20°C (after melting; black solid line), at 20°C (red solid line), at 95°C (green solid line)

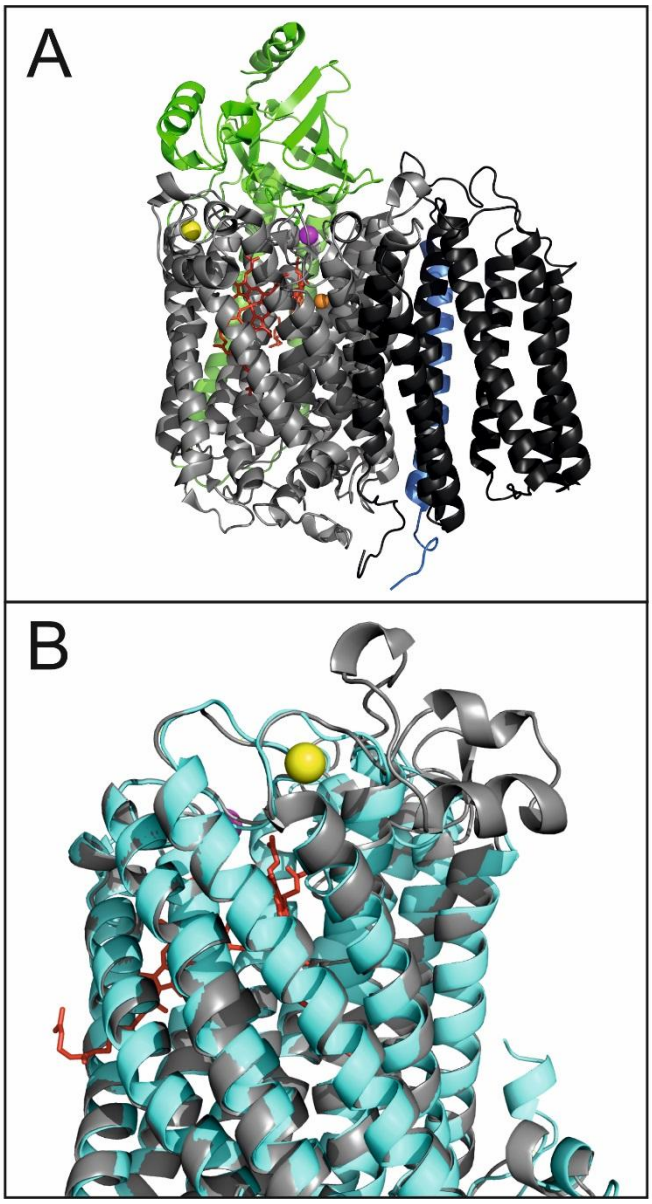


Figure S4: Calcium-binding in cytochrome *c* oxidase subunit COXA

A. Structure of cytochrome c oxidase from Rhodobactor sphaeroides (PBD 1M56) [74], showing subunits COXA (light grey), COXB (green), COXC (dark grey), COXD (blue). Cofactors shown are Ca^{2+} (yellow sphere), Cu^{2+} (orange sphere), Mg^{2+} (magenta sphere) and haem (red sticks); **B.** Predicted protein structure of *Anabaena* sp. PCC 7120 COXA (alr0951; cyan) aligned with *R. sphaeroides* COXA subunit (grey; same as in **A**), to highlight putative Ca^{2+}-binding site in *Anabaena* COXA.

Table S1 List of oligonucleotides used in the current study

Oligonucleotide	**Description or sequence**
cse-NdeI-S	5'-CCACTCCCATATGGCAACCGAGCAAGAGCTTCAA-3'
cse-EcoRI-AS	5'-GCCGCTGCCAGCGCTGAATTCCTAGGTTAAATTACTTGCTTTCTT-3'
EF1	5'-GCCAAAATAAATATCAGTAAA-3'
EF4	5'-AGCTGCGGTTGATTATTTGTATTA-3'
*rpoA*_qPCR-S	5'-CAACTCTCTGTACGGGCCTA-3'
*rpoA*_qPCR-AS	5'-GCTTCTTTCTTGGGGTAAGG-3'
*cse*_qPCR-S	5'-TCAAGACGGCAAAATCTCCA-3'
*cse*_qPCR-AS	5'-CTGTACTAGCAACTCTTGGGG-3'

***Conflict of Interest**

Click here to download Conflict of Interest: COI disclosure.pdf

<u>Publication 5:</u> in Scientific Reports; 9(1):18985.
doi: 10.1038/s41598-019-55495-y.

Tuning the *in vitro* sensing and signaling properties of cyanobacterial PII protein by mutation of key residues

Authors: Khaled A. Selim[a,1], Michael Haffner[a], Björn Watzer[a], Karl Forchhammer[a,1]

Affiliations:
[a] Interfaculty Institute for Microbiology and Infection Medicine, Organismic Interactions Department, Tübingen University, Auf der Morgenstelle 28, 72076 Tübingen, Germany.

[1]To whom correspondence may be addressed. E-mail: KAS: Khaled.selim@uni-tuebingen.de; KF: karl.forchhammer@uni-tuebingen.de

<u>Authors E-mails list:</u>

K.A. Selim: Khaled.selim@uni-tuebingen.de
M. Haffner: michael.haffner@student.uni-tuebingen.de
B. Watzer: Bjoern.Watzer@student.uni-tuebingen.de
K. Forchhammer: karl.forchhammer@uni-tuebingen.de

Abstract (249/250 word)

PII proteins are ancient signaling transduction proteins, widespread among all domains of the life from bacteria to Eukaryotic-plants. PII proteins regulate the central cellular metabolism by measuring and integrating the current carbon/nitrogen/energy status of the cell. So far, it's known that PII proteins exert their signaling function through binding ATP, ADP, and 2-oxogluterate. In cyanobacteria, PII regulates the nitrogen metabolism by interacting with: 1) N-acetyl-glutamate kinase (NAGK), the key enzyme of arginine biosynthesis, and 2) PII-interacting protein-X (PipX), the co-activator of the global nitrogen-trascription factor NtcA. The binding of PII to NAGK increases NAG/NAG-P turnover and prevents arginine-feedback inhibition to NAGK. The binding of PII to PipX protein leads to a deactivation of the NtcA regulon. To figure out the key residues for active cellular function of PII protein, we created different point mutations in cyanobaterial *Synechococcus elongatus* PCC 7942 PII protein. A variety of *in vitro* approaches were applied to study the influence of K58N and I86N point mutations on PII mediated NAGK and PipX interactions. Here, we show that the Lys58 is a key residue for mediating PII interactions. The K58N mutation causes the loss of 2-oxogluterate binding and a strong negative effect on ADP, NAGK and PipX binding. The replacement of the nearby Leu56 by Lys again can partially compensate the loss of K58. The interaction analysis using SPR revealed that I86N variant form a very strong complexes with NAGK and PipX *in vitro*, suggesting a huge impact of PII (I86N) on nitrogen metabolism *in vivo*.

Key words:
Cell signaling PII protein,
PII-interacting protein X (PipX),
N-acetyl-glutamate kinase (NAGK),
PII (K58N),
PII (I86N).

Introduction

The PII signal-transduction superfamily is widely spread in all domains of the life, representing one of the largest and most ancient families of signaling proteins in nature. Signaling proteins of the PII superfamily are characterized by their highly conserved trimeric structure, consisting of a triangular core of β-sheets from the ferredoxin-like fold of the three subunits (Forchhammer & Lüddecke 2016, Selim et al. 2018). Canonical PII proteins are involved in the regulation of various carbon and nitrogen metabolism related processes (Forchhammer 2008, Huergo et al. 2013, Forchhammer & Lüddecke 2016). For this purpose, PII acts as an energy/carbon/nitrogen sensor: binding of metabolites as well as post-translational modifications enable PII to integrate different signals (Fokina et al. 2010a, Zeth et al. 2014, Forchhammer 2008, Forchhammer & Lüddecke 2016). Based on this input signals, PII triggers a specific output signal by interacting with target proteins including regulatory proteins, transcription factors, enzymes of carbon/nitrogen metabolism or transporters (Huergo et al. 2013, Forchhammer & Lüddecke 2016). The competitive binding of ATP or ADP and synergistic binding of 2-oxogluterate (2-OG) allows PII to estimate the current energy and nitrogen/carbon status of the cell (Fokina et al. 2010a&b, Fokina et al. 2011, Zeth et al. 2014, Forchhammer & Lüddecke 2016, Lapina et al. 2018). Canonical PII proteins are highly conserved homo-trimeric proteins with three characteristic loop regions (T-, B- and C-loops). These loop regions are located near the inter-subunit clefts and play a major role in protein-protein interactions and ligand binding (Fokina et al. 2010a&b, Zeth et al. 2014, Forchhammer & Lüddecke 2016).

In cyanobacteria and plants, the major interaction partners of PII described so far are the controlling enzyme or arginine synthesis, N-acetyl-L-glutamate kinase (NAGK), the BCCP-subunit of Acetyl-CoA Carboxylase (ACCase) and restricted to cyanobacteria only, the transcriptional co-activator PipX (Llácer et al. 2007, 2010, Forchhammer & Lüddecke 2016, Hauf et al. 2016, Lapina et al. 2018). PipX acts as a transcriptional co-activator of the master transcription factor of nitrogen regulated genes, NtcA (Llácer et al. 2007, Llácer et al. 2010, Espinosa et al. 2006; 2007; 2014, Heinrich et al. 2004, Beez et al. 2009, Fokina et al. 2010b, Fokina et al. 2011, Forchhammer & Lüddecke 2016). Under sufficient nitrogen supply, indicated by a low 2-OG levels, PII forms an ATP-dependent activating complex with NAGK and complexes, preferentially in the ADP-bound state, with PipX, causing a decreased in NtcA-PipX complex formation (Fokina et al., 2011). The PII-NAGK complex formation increases the catalytically efficiency of NAGK and decrease the feedback inhibitory effect of arginine to the NAGK (Heinrich et al. 2004, Beez et al. 2009). Under conditions of high 2-OG levels (poor nitrogen supply), the ATP-dependent binding of 2-OG to PII causes strong conformational

changes in T-loop, which in turn impairs the interaction of PII proteins with both NAGK and PipX (Fokina et al., 2010). The binding of PII in the ATP-bound state to BCCP-subunit of ACCase inhibits the ACCase activity (Feria-Bourrellier et al. 2010, Gerhardt et al. 2015, Hauf et al. 2016). In all these cases, the PII proteins exert their regulator function through binding the small effector molecules ATP/ADP and 2-OG, which induces conformational changes on the flexible surface exposed T-loop, the PII's major protein-interaction structure (Zeth et al. 2014).

Previously, we showed that replacement of Ile86 with Asn causes a folding of the PII T-loop in a bent conformation similar to the T-loop conformation in the PII-NAGK complex (Fokina et al. 2010b). As a consequence of the I86N mutation, the PII (I86N) variant can bind constitutively to NAGK *in vitro* and *in vivo*, leading to an accumulation of arginine and cyanophycin (Fokina et al. 2010b and Watzer et al. 2015). Lys58 is known to interact with B-loops through direct interaction with G87, and with the T-loop via salt bridge with E44 when PII is complexed with NAGK (Llácer et al. 2007, Truan et al. 2010, Fokina et al. 2010a&b). To gain deeper insights in the signaling mechanism of PII from cyanobacterium *Synechococcus elongatus* PCC 7942, we investigated the influence of PII-I86N mutation on the interaction with PipX and furthermore, checked the importance of Lys58 for PII functions, in particular, for the PII-interaction with NAGK and PipX. We used PII protein from cyanobacterial model organism (Fokina et al. 2010a&b, Fokina et al. 2011).

Results

K58 is crucial for ligand binding.

The PII-NAGK complex formation requires a compact conformation of the protruding PII T- loop. The T-loop is inserted in the interdomain cleft of the NAGK subunits for anchoring PII on NAGK (Llácer et al. 2007). Residue Lys58 stabilizes the tight folding of the T-loop in the PII-NAGK complex through a salt bridge with Glu44 of PII T-loop (Llácer et al. 2007, Fokina et al. 2010a&b, Fokina et al. 2011). When the levels of 2-OG increase, 2-OG can bind to the PII-Mg-ATP complex, involving a salt-bridge of the 2-OG C5-carboxyl-group with the positively charged K58 side chain (Llácer et al. 2007 & Fokina et al., 2010a&b). When PII is in the PII-NAGK complex, binding of 2-OG requires that the intramolecular K58-E44 salt breaks, which should open the T-loop and dissociate the PII-NAGK complex (Llácer et al. 2007 & Fokina et al., 2010a&b, Fokina et al. 2011). To directly assess the importance of PII Lys58 for NAGK complex formation, we created a *Se*PII (K58N) variant, where Lys58 was replaced by Asp, and therefore the K58-E44 salt bridge cannot form. Since K58 is a ligand for 2-OG binding, the binding properties of *Se*PII (K58N) variant in comparison to the *Se*PII (WT) protein were analyzed by

isothermal titration calorimetry (ITC) (Fig. 1). As expected, the PII variant K58N lost the ability to bind 2-OG (Supplementary Fig. S1), whereas ATP was still able to bind (Fig. 1A), albeit weaker than the *Se*PII (WT) protein (Table 1). The binding of ADP was even more strongly inhibited (Fig. 1B), yielding only very weak isotherm signals. This agrees with the fact that the K58 residue is crucial for establishing a stable T-loop conformation of the ADP-complex through hydrogen-bonding interaction with the Q39 side chain (Truan et al. 2014). Structural examination of the *Se*PII-ADP structure shows that the side chain of the neighboring residue Leu56 also approaches the Q39 side chain, suggesting that replacement of L56 by K might be able to compensate the K58N mutation (PDB: 2XUL and 2XBP). To test this hypothesis, we created double point mutation variant *Se*PII (K58N/L56K). The *Se*PII (K58N/L56K) variant was tested for its binding properties using ITC. This PII variant was still unable to bind 2-OG (Supplementary Fig. S1) but enhanced the binding affinity toward ATP in comparison to the *Se*PII (K58N) variant (Fig. 1, compare A with C). Especially, the first biding site was occupied with very high affinity. The ADP binding events to *Se*PII (K58N/L56K) variant induced a strong isotherm, and in particular, the affinity for the second binding site was strongly enhanced (Table 1, Fig. 1, compare D with B). As can be deduced from the binding isotherms, the binding enthalpy for ATP and ADP binding to *Se*PII (K58N/L56K) variant was stronger than to *Se*PII (K58N) variant (Fig. 1, compare A and B with C and D). These results indicate that a relocation of the Lys-residue by two amino acid positions doesn't rescue 2-OG binding but has a positive effect on adenyl nucleotide binding, in agreement with the fact that K58 is a direct ligand to 2-OG but is only indirectly involved in adenyl-nucleotide binding through affecting the T-loop conformation.

PII variants K58N and K58N/L56K are still able to activate NAGK.

Next, we wanted to investigate the ability of *Se*PII variants (K58N and K58N/L56K) to activate NAGK. The PII-based activation of NAGK was tested using a coupled enzyme assay (Beez et al. 2009) with recombinant NAGK proteins deriving from strains *S. elongatus* PCC7942 (*Se*NAGK) and *Synechocystis* sp. PCC 6803 (*Sc*NAGK). The kinetic activities of the *Se*NAGK (listed in Table 2) was determined with NAG (N-acetyl-glutamate) as a variable substrate in presence or absence of different PII protein variants. With *Se*NAGK, both PII variants were able to activate *Se*NAGK (Fig. 2A), but weaker than the wild-type variant *Se*PII (WT). With *Sc*NAGK, the *Se*PII (K58N) variant was not able to activate *Sc*NAGK (Supplementary Fig. S2) but interestingly, the *Se*PII (K58N/L56K) variant was able partially to activate *Sc*NAGK (Supplementary Fig. S2). This indicates that the interaction of PII with the non-cognate NAGK is

more susceptible to mutations that affect the T-loop conformation (K58N) and that relocation of Lys58 to position 56 allows PII to adopt a compensatory conformation for NAGK interaction.

Tight complex formation of PII-NAGK is required to relieve arginine feedback inhibition.

The relief of NAGK from arginine feedback inhibition through PII-NAGK complex formation is the rate-limiting step for the metabolic switch of the arginine biosynthetic pathway (Beez et al. 2009, Lapina et al. 2018). To examine whether the different PII variants would relief NAGK from arginine inhibition, we assessed *Se*NAGK activity at fixed concentration of NAG (40mM) in presence of different concentrations of arginine with and without different PII protein variants. With *Se*NAGK, the arginine feedback inhibition occurred with a half minimal inhibitory concentration (IC_{50}) of 11 µM (Fig. 2B). As expected, the arginine inhibitory effect on *Se*NAGK was completely released in presence of *Se*PII (WT) protein (Fig. 2B). Both *Se*PII variants (K58N and K58N/L56K) were able to activate *Se*NAGK, but the complexes were still sensitive to arginine feedback inhibition (IC_{50} of 14.2 and 15.7, respectively), which is slight less sensitive towards arginine than NAGK alone (Fig. 2B). Remarkably, the double point mutation variant activated *Se*NAGK stronger and showed less sensitivity to arginine inhibition than the single point mutation (K58N) (Fig. 2B), implying that replacement of Leu56 with Lys could restore partially the loss function of Lys58 in stabilizing PII-NAGK complex.

Determination of the kinetic activity of the NAGK with NAG as a variable substrate in the presence of 11 µM arginine (which corresponds to the IC_{50} of NAGK alone) with and without different PII variants discerned the two mutant PII variants. In principle, both PII variants were able to activate *Se*NAGK (Fig. 2C), but weaker than the *Se*PII (WT). However, the *Se*PII (K58N/L56K) variant was much more efficient than the *Se*PII (K58N) variant. The kinetic data for the *Se*PII (K58N) variant could not be fitted to Michaelis-Menten kinetics anymore. Similar results were obtained using the *Sc*NAGK enzyme (Supplementary Fig. S2).

Next, we determined the response of the various PII-NAGK complexes to 2-OG. As shown in Fig. (2D), the addition of 2-OG had a negligible influence on both variants of PII proteins conforming the inability of K58 mutant to bind 2-OG effectively (Fig. 2D). In contrast, the *Se*PII (WT) protein showed the typical inhibitory response in presence of 2-OG (Fig. 2D).

Mutation of PII K58 impairs PII–NAGK complex formation.

Form the previous enzymatic assays, it turned out that both PII variants were still able to interact with NAGK, but this interaction seemed to be impaired to different extent, with the (K58N/L56K) variant partially compensating the impairment of the K58N variant. To further confirm this

hypothesis, we first tried to isolate PII-NAGK complexes of different variants of PII (K58N) and (K58N/L56K) using analytical gel-filtration coupled with multi angle light scattering (MALS) (Fig. 3A). Wild-type *Se*PII protein forms a stable complex with *Se*NAGK with a molar-mass of 275.4 kDa, corresponding to two-PII trimers (each trimer of ≈ 40.8 kDa) sandwiching one hexameric *Se*NAGK (≈ 194 kDa) (Llácer et al. 2007). In this experimental setting, both proteins eluted together from the gel-filtration column with an apparent mass according to MALS analysis of 255.9 kDa (Fig. 3A), which agrees with an almost complete NAGK-PII complex. However, with both variants, we were unable to detect any indications of a *Se*NAGK-PII complex, since *Se*NAGK eluted as free hexamer with an apparent mass of 193.8 kDa (Fig. 3A). In agreement, the corresponding fractions as analyzed by SDS-PAGE contained *Se*NAGK without any traces of PII (Fig. 3B). This experiment confirmed that the physical interaction between *Se*NAGK and different variants of PII protein is weak and the weak complexes dissociate during the gel-filtration.

Additionally, we assessed whether binding of *Se*PII (K58N) and (K58N/L56K) variants to *Se*NAGK could be measured with the more sensitive method using the surface plasmon resonance (SPR) spectroscopy (Fig. 3C). The His-tagged *Se*NAGK was immobilized on a Ni-NTA sensor chip and PII variants (K58N) and (K58N/L56K), as well as *Se*PII (WT) protein, were injected sequentially. No interaction between *Se*NAGK and PII variants (K58N) and (K58N/L56K) could be detected.

K58N mutation influences negatively also on the PII-PipX complex formation.

Another major interacting partner of the PII signaling network in cyanobacteria is PipX (Forcada-Nadal et al. 2018). In the absence of 2-OG, three PipX monomers can bind to one PII trimer (Llácer et al. 2010, Fokina et al. 2011). The crystal structure of the PII-PipX complex revealed that the T-loop of PII acts as an antenna that attracts the PipX (Llácer et al. 2010, Zhao et al. 2010). Unlike the bent conformation of the T-loop in NAGK-PII complex, the T-loop in the PipX-PII complex is in an extended conformation, resembling PII in complex ADP (Zeth et al., 2014). PII in the Mg-ATP-2-OG complex is unable to bind PipX. Since the PII (K58N) mutation impaired ADP binding (Fig. 1), we assumed that the K58N variant may be affected in PII-PipX complex formation. We analyzed the formation of PII-PipX complexes by SPR spectroscopy, using an indirect assay, as described previously (Espinosa et al. 2006, Fokina et al. 2011). When PipX is incubated with PII prior the injection on the sensor chip, PII-PipX complex formation increases the binding of His_6-PipX to the sensor chip due to mass increase and PipX trimerization (Fig. 4A). In absence of the effector molecules, the PII (K58N) variant was not able to increase the

PipX binding, while the PII (K58N/L56K) was able to partially restore the binding of His_6-PipX to the sensor chip, as compared to PII (WT) (Fig. 4A). In presence of 3mM ADP, the PII (K58N) variant regained the ability to at least partially bind PipX (Fig. 4B). Again, the PII (K58N/L56K) variant showed a stronger interaction with PipX than the PII (K58N) variant but still weaker than PII (WT) (Fig. 4B). To compare the effect of PII variants on the dissociation rate of the complexes from the sensor chip (which is a good indicator of the efficiency of PII-PipX interaction), the dissociation curves were normalized to the RUs at the end of the injection phase (taken as 100 %) (Fig. 4C). In presence of ADP, the PipX-PII (K58N/L56K) variant complex dissociates slowly in comparable way to the PipX-PII (WT) complex, while the PipX-PII (K58N) complex dissociates faster (Fig. 4*C*), confirming again that L56K could compensate the loss of Lys58. In presence of 1mM ATP, PII (WT) was able to bind to PipX with less efficiency as in presence of ADP, in agreement with previous reports (Espinosa et al. 2006, Fokina et al. 2011), while both PII variants (K58N) and (K58N/L56K) were not able to efficiently interact with PipX (Fig. 4D).

The T-loop bent conformation facilitates PII-PipX interaction.

Previously, we reported a NAGK hyper-activating variant PII (I86N) (Fokina et al. 2010b, Watzer et al. 2015). The I86N substitution causes the T-loop of PII to adopt a compact conformation through formation of hydrogen bond between the backbone oxygen of T43 and the amido group of N86 resulting in a contraction of the T-loop (Fokina et al. 2010b). As consequence, the PII (I86N) variant binds constitutively to NAGK. However, the interaction of PII (I86N) variant with PipX has not been analyzed before. When PII proteins were pre-incubated with PipX in the absence of effector molecules, the P_{II} (I86N) variant lead to a much stronger increase of binding of PipX to the Ni-NTA sensor chip surface and to decreased dissociation as compared to PII (WT) (Fig. 5A). To compare the effect of PII proteins on the dissociation of the complexes from the sensor chip, the dissociation curves were normalized to the RUs at the end of the injection phase (taken as 100%) (Fig. 5B). The percent RUs remaining bound to the chip after 400 s of dissociation were then taken as a proxy for PII-PipX interaction and used to quantify the effect of different effector molecules (Fig. 5B-E).

In the following experiments, different effector molecules were tested on PII-PipX complex stability. As expected, a strong positive effect of ADP on the interaction of PII (WT) with PipX, was obtained (Fokina et al. 2011), (Fig. 5D & E). By contrast, binding of the PII (I86N) variant to PipX was negatively affected by ADP (Fig. 5 C & E). ATP showed for the PII (WT) protein a slightly lower stability of the complex than in the ADP-complexed state whereas the PII (I86N)

variant interacted stronger with PipX in the ATP state than with ADP (Fig. 5 C, D & E). As expected, 2-OG in presence of ATP impaires P_{II}-PipX complex formation (Espinosa et al. 2006, Fokina et al. 2011) (Fig. 5 D). An inhibitory effect of ATP and 2-OG on the PII-PipX complex was also visible with the PII (I86N) variant, however not as strong as with PII (WT) protein (Fig. 5 C, D & E). This agrees with the impaired ability of the PII (I86N) variant to bind 2-OG (Fokina et al. 2010b). Taken together, these data demonstrated that the PII (I86N) variant is very efficient in complex formation with PipX and that this binding does not require positive stimulation by ADP as is the case with wild-type PII.

Discussion

Previous structural analysis of PII proteins in various complexed states showed that Lys58 plays an important role in the function of PII (Fokina et al. 2010a, Truan et al. 2010, Truan et al. 2014, Zeth et al. 2014). However, the K58 is not involved in direct interactions with ADP or ATP, but it rather seems to be important to anchor the B-loop via a H-bond to B-loop residue G87 (PDBs: 2XUL, 3MHY, and 2XBP). The ADP binding events induces conformational changes within the surface exposed T-loop enabling K58 to interact with Q39 at the base of the T-loop. Therefore, K58 can serve as well to anchor the T-loop in the ADP bound conformations via hydrogen bond to Q39 (PDB: 4CNZ), which characterize/facilitate the preferred ADP bound state (Truan et al. 2014, Zeth et al. 2014). Further, it anchors also the T-loop in the bent conformation of the PII-NAGK complex via a salt-bridge to E44 (PDBs: 2V5H & 2XBP) (Llácer et al. 2007, Truan et al. 2010 & 2014, Fokina et al. 2010a&b). Its notable that the K58-network interactions drags the basal part of the T-loop (around Q39) close to the B loop (I86/G89) (Zeth et al. 2014), causing subtle conformational changes within the T-loop towards the interacting partners.

Remarkably, the 2-OG binding is coordinated mainly by the highly conserved residues of PII proteins K58 and Q39 (Fokina et al. 2010a, Truan et al. 2010). Upon 2-OG binding, the K58-E44 interaction is broken and replaced by a new salt bridge interaction between K58 and 2-OG (PDBs: 2XUL & 3MHY). Elegantly, the C5 carboxyl group of 2-OG quite precisely displaces the E44 carboxyl group, such that K58 does not need to change its conformation. Therefore, we expected the K58N variant to be defective in both 2-OG binding and in anchoring the T-loop in the NAGK-bound conformation.

Complex formation of PII with NAGK follows a two-step mechanism (Fokina et al. 2010b): first, an encounter complex involving the B-loop of PII is formed. This involves a contact between R233 of NAGK and E85 of PII. This interaction breaks a salt bridge between E85 and R47 in the T-loop of PII, which is now free and can adopt a bended structure. The T-loop bended

confirmation is stabilized by a new salt bridge between K58 and T-loop residue E44. In the second step, the bended T-loop deeply inserts into the NAGK clefts to form the tight complex. Apparently, the mutation of K58 destabilize the PII structure, but the T-loop can still insert into the NAGK clefts with the help of the other hydrophobic interactions between I229, I253, and A257 of NAGK and F11 and T83 of PII (Llácer et al. 2007), however in presence of arginine the complex became more sensitive to Arg inhibition and dissociates.

Our *in vitro* assays confirmed that the K58 is a key residual for proper signaling function of PII via affecting on: 1) sensory properties of ATP/ADP/2-OG and 2) the interactions with NAGK and PipX. The PII (K58N) variant is still able to bind ATP but lost the ability to bind 2-OG and strongly influence ADP binding negatively. The loss of 2-OG sensing is expected, since K58 contribute binds directly 2-OG. Constant with our results, previous results demonstrated that the PII (K58M) was not able to bind to 2-OG (Fokina et al. 2010a). However, we believe strongly that the defect in anchoring the T-loop (and maybe B-loop) is most likely the cause for the changing the affinities of the adenyl nucleotides (ATP and ADP). Thus, it seems that anchoring the T-loop is crucial for nucleotides binding. Apparently, the mutation of K58 destabilizes the PII T-loop structure, since the NAGK and PipX interactions which are mediated by the T-loop are negatively influenced by K58N mutation. In case of NAGK, the T-loop of PII (K58N) variant can still activate the NAGK, however in presence of arginine the complex became more sensitive and dissociated. Therefore, we concluded that the tight complex formation between PII and NAGK is required to relief NAGK from Arg feedback inhibition. The complementation of K58N with a second mutation in the nearby L56 to Lys again can restore partially the ability of PII to interact with NAGK and PipX and can be explained by the ability of L56K to approach again E44, which can stabilize partially again the T-loop.

In cyanobacteria, the regulation of the nitrogen metabolism mainly depends on the signal-transduction protein PII, the transcription factor NtcA and PipX Protein (the NtcA or PII interacting partner). The crystal structure of PII (I86N) (PDB: 2XBP) showed an almost identical backbone as PII (WT). However, the T-loop adopts a compact conformation, which is a structural mimic of PII in NAGK complex (Fokina et al. 2010b). Since, the PII (I86N) variant strongly activates NAGK *in vitro* and *in vivo*, causing huge accumulation of Arg and cyanophycin (Fokina et al. 2010b, Watzer et al. 2015). We previously suggested that this variant is unable to interact with other PII interaction partners. Therefore, it was interesting to characterize the response of that variant regarding PipX protein, which in return could influence the NtcA regulon. The SPR spectroscopy confirmed that the PII (I86N) maintains the ability to bind PipX. Moreover, in absence of effector molecules, the PII (I86N) variant showed stronger interaction

with PipX than PII (WT). Next, we tested the influence of effector molecules to the complex formation. The PII (I86N) variant formed complexes more stable than the PII (WT) under all tested condition. Furthermore, the PII (I86N) variant did not require positive stimulation by ADP to form a sable complex as it is the case of PII (WT). It was formally shown that the PII (I86N) variant did not respond to 2-OG in complex with NAGK (Fokina et al. 2010b). An inhibitory effect of ATP and 2-OG on the PII-PipX complex is also visible with the PII (I86N) variant, although not as strong as in the case of the PII (WT) protein. Taken together, the interaction of PII (I86N) and PipX is deregulated, but the complex stability is generally higher compared to PII (WT) with PipX. Due to the stronger sequestration of PipX by PII (I86N), we speculate that the PII (I86N) might tune down the 2-OG mediated NtcA response *in vivo*. Therefore, we suggested a delayed nitrogen starvation response in a *Synechocystis* sp. harboring the PII (I86N) variant. As a future perspective, a proteomic or transcriptomic analysis would be informative to characterize the *in vivo* influence of PII (I86N) on the NtcA regulon, in response to nitrogen limitation conditions as the *Synechocystis* sp. BW86 strain carrying the PII (I86N) variant is available (Watzer et al. 2015).

Altogether, our study reveals new insights into the signaling function of PII protein and sheds some lights on the plasticity of PII body, due to for example, the ability of PII (K58N/L56K) to compensate partially the loss of K58 confirms the inherited flexibility of PII body. This flexibility of PII body could be used in protein engineering to design new proteins or sensors with new desired characters. Notably, the PII protein was used previously to engineer PII variant (I86N) to sense citrate as a new effector molecule (Zeth et al. 2012). Also, in several studies demonstrated engineering of 2-OG sensor proteins based on trimeric architecture of PII protein (Lüddecke et al. 2017, Chen et al. 2018).

Material and methods

Cloning, overexpression and purification of recombinant PII variants, NAGK, and PipX proteins: The recombinant strepII-tagged PII proteins (wildtype and I86N) encoded by *glnB* gene and *glnB*(I86N) from *S. elongatus* were cloned into the Strep-tag fusion vector pASK-IBA3 to generate PII(WT)-pASK-IBA3 and PII(I86N)-pASK-IBA3 plasmids as described previously (Heinrich et al. 2004, Beez et al. 2009, Fokina et al. 2010b). The strep-tagged PII variants (K58N) and (K58N/L56K) were generated by amplifying the whole PII(WT)-pASK-IBA3 plasmid using site-directed mutagenesis primers. The PII(K58N)-pASK-IBA3 plasmid was generated using 1848_Fw: GGTTGAGTTTTTGCAAAATCTGAAGCTCGAG and 1849_Rv: GTGTATTCCGAGCCGCGATAGC primers, while the PII(K58N/L56K)-pASK-IBA3 plasmid was generated using 1850_Fw: GGTTGAGTTTAAGCAAAATCTGAAGCTCG and 1849_Rv primers. The generated plasmids were

verified by sequencing. The pASK-IBA3 based plasmids were transformed and overexpressed in PII deficient *E. coli* strain RB9060 (Lapina et al. 2018). The recombinant strep tagged PII proteins were purified using affinity chromatography as described previously (Heinrich et al. 2004, Beez et al. 2009, Fokina et al. 2010b, Selim et al. 2018). The His_6-tagged NAGK and PipX were overexpressed in *E. coli* BL21(DE3) and purified via affinity chromatography using HisTrap column (GE Healthcare) as described previously (Heinrich et al. 2004, Espinosa et al. 2006, Beez et al. 2009, Lapina et al. 2018).

Surface plasmon resonance (SPR) spectroscopy for determination of PII-NAGK and PII-PipX complex formation: The SPR experiments were performed using a BIAcore X biosensor system (GE Healthcare) as described previously (Beez et al. 2009, Fokina et al. 2010b, Fokina et al. 2011) at 25℃ with HBS-Buffer (10mM Hepes pH 7.5, 150mM NaCl, 1mM $MgCl_2$ and 0.005% (w/v) Nonidet P40). For a typical assay, 100-1000 nM Strep-PII (WT or variants) was preincubated in the absence and in presence of effectors (1mM of ADP; ATP; ATP/2-OG) for 10min at RT. To assess the ability of PII variants (K58N, K58N/L56K, and I86N) to complex with PipX, an indirect assay was used as described previously (Espinosa et al. 2006, Fokina et al. 2011). Briefly, the assay measures the interaction of His_6-PipX with a Ni-NTA sensor chip in the presence or absence of PII. Upon PII interaction, monomeric PipX protein trimerizes at the PII interaction surface, which leads to a stabilization of the His_6-PipX interaction with the Ni-NTA surface. Furthermore, the additional mass of the PII protein leads to an increased SPR signal. For PipX assays, 500 nM His_6-PipX was added to the preincubated Strep-PII and further incubated for 10min at room temperature. The PII-PipX mixture was injected on a Ni^{2+}-loaded NTA chip to Flow Cell 2 (FC2), while FC1 (without Ni^{2+}) was used as a background control. As control, His_6-PipX was injected to the chip in absence of PII. Injection phase takes 200s (flow rate 15 µl/s); followed by dissociation at a flow rate of 15 µl/s. ΔRUs at the end of injection phase, were normalized to 100 % to compare the dissociations. ΔRUs after 400s of dissociation were used as a relative value of complex stability. To evaluate the extend of PII variants PII variants (K58N and K58N/L56K) to interact with His_6-NAGK, standard SPR experiments were performed as described previously (Beez et al. 2009, Fokina et al. 2010b). Beefily, His_6-NAGK was immobilized on a Ni^{2+}-loaded NTA chip to Flow Cell 2 (FC2), then Strep-PII proteins was injected on FC1 and FC2, the difference in response between the FC2-FC1 represents the specific interaction between PII proteins and NAGK. To load fresh proteins on to the NTA sensor chip, bound proteins were removed by injecting 25 µl of 0.4 M EDTA pH 7.5 (flow rate 15 µl/s). Subsequently, 10 µl of 5 mM Ni_2SO_4 (flow rate 10 µl/s) was injected to FC2.

Isothermal titration calorimetry (ITC): ITC experiments were performed using VP-ITC microcalorimeter (MicroCal) in a buffer composed of 50 mM Tris-HCl (pH 8.0) and 200 mM NaCl at 20℃, according to (Lapina et al. 2018). The row calorimetric data were fitted according to three-sequential binding sites models using MicroCal Origin software (Lapina et al. 2018, Selim et al. 2018).

Coupled NAGK activity assay: To assess the influence of PII variants (K58N and K58N/L56K) on the NAGK activity, we used our standard NAGK-coupled assay in which the ADP production is coupled with the oxidation of NADH using the auxiliary enzymes pyruvate kinas and lactate dehydrogenase as described previously (Beez et al. 2009, Lapina et al. 2018).

Size exclusion chromatography coupled to multiangle angles light scattering (SEC-MALS) analysis: SEC-MALS was carried out using an Äkta chromatography (GE Healthcare) commented to MALS system (Wyatt Technology Corp.) at room temperature, according to (Selim et al. 2019). SEC-MALS experiments were done on a 24 ml geometric Superose 6 Increase 10/300 GL column (GE Healthcare) to which an Optilab T-rEX refractometer and a miniDawn Treos systems (Wyatt) were attached. The runs were performed with a 0.5 ml/min flowrate after equilibration of the column with the running buffer (100 mM Tris-HCl/pH 7.8, 300 mM NaCl, 2 mM $MgCl_2$, 2% glycerol). The PII proteins in excess were mixed with NAGK (3:1) for 10 min at room temp, then 100 µl of the mixes were injected to investigate the ability to isolate PII-NAGK complex. The elution volume was plotted against the UV signal and the molecular mass derived from the light scattering data. The data analysis and molecular weight calculations were done using ASTRA software (Wyatt) (Selim et al. 2019). The elution fractions were collected and subjected to SDS-PAGE for analysis.

Acknowledgements

This work was supported by DFG to K.F. (Fo195/9-2, Fo195/13-1). K.A.S. received a scholarship from the Deutsche Akademische Austauschdienst (DAAD).

Author contributions

K.A.S. and K.F. designed the study and wrote the manuscript. K.A.S. generated and purified PII (K58N and K58N/L56K) variants and performed ITC, SPR and SEC-MALS experiments. M.H. characterized NAGK enzyme activity. B.W. characterized PII (I86N) variant. All authors analyzed the results and approved the final version of the manuscript.

Additional Information.

Supplementary information accompanies this paper at

Competing Interests: The authors declare that they have no competing interests.

References

1. Beez S, Fokina O, Herrmann C, Forchhammer K. N-acetyl-L-glutamate kinase (NAGK) from oxygenic phototrophs: P(II) signal transduction across domains of life reveals novel insights in NAGK control. J Mol Biol. 2009; 389(4):748-58.

2. Chen HL, Latifi A, Zhang CC, Bernard CS. Biosensors-Based In Vivo Quantification of 2-Oxoglutarate in Cyanobacteria and Proteobacteria. Life (Basel). 2018; 8(4).

3. Espinosa J, Forchhammer K, Burillo S, Contreras A. Interaction network in cyanobacterial nitrogen regulation: PipX, a protein that interacts in a 2-oxoglutarate dependent manner with PII and NtcA. Mol Microbiol. 2006 Jul;61(2):457-69.

4. Espinosa J, Forchhammer K, Contreras A. Role of the Synechococcus PCC 7942 nitrogen regulator protein PipX in NtcA-controlled processes. Microbiology. 2007 Mar;153(Pt 3):711-8.

5. Espinosa J, Rodríguez-Mateos F, Salinas P, Lanza VF, Dixon R, de la Cruz F, Contreras A. PipX, the coactivator of NtcA, is a global regulator in cyanobacteria. Proc Natl Acad Sci U S A. 2014; 111(23):E2423-30.

6. Feria Bourrellier AB, Valot B, Guillot A, Ambard-Bretteville F, Vidal J, Hodges M. Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate-regulated by interaction of PII with the biotin carboxyl carrier subunit. Proc Natl Acad Sci U S A. 2010; 107(1):502-7.

7. Fokina O, Chellamuthu VR, Forchhammer K, Zeth K. Mechanism of 2-oxoglutarate signaling by the Synechococcus elongatus PII signal transduction protein. Proc Natl Acad Sci U S A. 2010a; 107(46):19760-5.

8. Fokina O, Chellamuthu VR, Zeth K, Forchhammer K. A novel signal transduction protein P(II) variant from Synechococcus elongatus PCC 7942 indicates a two-step process for NAGK-P(II) complex formation. J Mol Biol. 2010b; 399(3):410-21.

9. Fokina O, Herrmann C, Forchhammer K. Signal-transduction protein P(II) from *Synechococcus elongatus* PCC 7942 senses low adenylate energy charge *in vitro*. Biochem J. 2011; 440(1):147-56.

10. Forcada-Nadal A, Llácer JL, Contreras A, Marco-Marín C, Rubio V. (2018) The PII-NAGK-PipX-NtcA Regulatory Axis of Cyanobacteria: A Tale of Changing Partners, Allosteric Effectors and Non-covalent Interactions. Front Mol Biosci. 5:91.

11. Forchhammer K, Lüddecke J[1]. Sensory properties of the PII signalling protein family. FEBS J. 2016; 283(3):425-37.

12. Gerhardt EC, et al. The bacterial signal transduction protein GlnB regulates the committed step in fatty acid biosynthesis by acting as a dissociable regulatory subunit of acetyl-CoA carboxylase. Mol Microbiol. 2015; 95(6):1025-35.

13. Hauf W, Schmid K, Gerhardt EC, Huergo LF, Forchhammer K. Interaction of the Nitrogen Regulatory Protein GlnB (P_{II}) with Biotin Carboxyl Carrier Protein (BCCP) Controls Acetyl-CoA Levels in the Cyanobacterium *Synechocystis* sp. PCC 6803. Front Microbiol. 2016; 7:1700.

14. Heinrich A, Maheswaran M, Ruppert U, Forchhammer K. The *Synechococcus elongatus* P signal transduction protein controls arginine synthesis by complex formation with N-acetyl-L-glutamate kinase. Mol Microbiol. 2004; 52(5):1303-14.

15. Huergo LF, Chandra G, Merrick M. 2013. P(II) signal transduction proteins: nitrogen regulation and beyond. *FEMS Microbiol Rev*. 37(2):251-283.

16. Lapina T, Selim KA, Forchhammer K, Ermilova E. The PII signaling protein from red algae represents an evolutionary link between cyanobacterial and Chloroplastida PII proteins. Sci Rep. 2018; 8(1):790.

17. Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K, Rubio V. Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. Proc Natl Acad Sci U S A. 2010 Aug 31;107(35):15397-402.

18. Llácer JL, et al. The crystal structure of the complex of PII and acetylglutamate kinase reveals how PII controls the storage of nitrogen as arginine. Proc Natl Acad Sci U S A. 2007; 104(45):17644-9.

19. Lüddecke J, Francois L, Spät P, Watzer B, Chilczuk T, Poschet G, Hell R, Radlwimmer B, Forchhammer K. P_{II} Protein-Derived FRET Sensors for Quantification and Live-Cell Imaging of 2-Oxoglutarate. Sci Rep. 2017; 7(1):1437.

20. Selim KA, Haase F, Hartmann MD, Hagemann M, Forchhammer K[1]. P_{II}-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. Proc Natl Acad Sci U S A. 2018; 115(21):E4861-E4869.

21. Selim KA, Lapina T, Forchhammer K, Ermilova E. Evolution of non-phototrophic metabolism in *Polytomella parva* alga involves unique adaptation of N-acetyl-L-glutamate kinase and PII sensing. Submitted to New phycologist (2019).

22. Truan, D., Bjelic, S., Li, X. D., Winkler, F. K. Structure and thermodynamics of effector molecule binding to the nitrogen signal transduction PII protein GlnZ from *Azospirillum brasilense*. *J. Mol. Biol*. 426, 2783-2799 (2014).

23. Truan, D., Huergo, L. F., Chubatsu, L. S., Merrick, M., Li, X. D., and Winkler, F. K. (2010) A new PII protein structure identifies the 2-oxoglutarate binding site. *J. Mol. Biol.* 400, 531–539.

24. Watzer B, Engelbrecht A, Hauf W, Stahl M, Maldener I, Forchhammer K. 2015. Metabolic pathway engineering using the central signal processor PII. *Microb Cell Fact*. 14:192.

25. Zeth K, Fokina O, Forchhammer K. An engineered PII protein variant that senses a novel ligand: atomic resolution structure of the complex with citrate. Acta Crystallogr D Biol Crystallogr. 2012; 68:901-8.

26. Zeth K, Fokina O, Forchhammer K. Structural basis and target-specific modulation of ADP sensing by the Synechococcus elongatus PII signaling protein. J Biol Chem. 2014; 289(13):8960-72.

27. Zhao MX, Jiang YL, Xu BY, Chen Y, Zhang CC, Zhou CZ. Crystal structure of the cyanobacterial signal transduction protein PII in complex with PipX. J Mol Biol. 2010 Sep 24;402(3):552-9.

Figures

Fig. 1. ITC

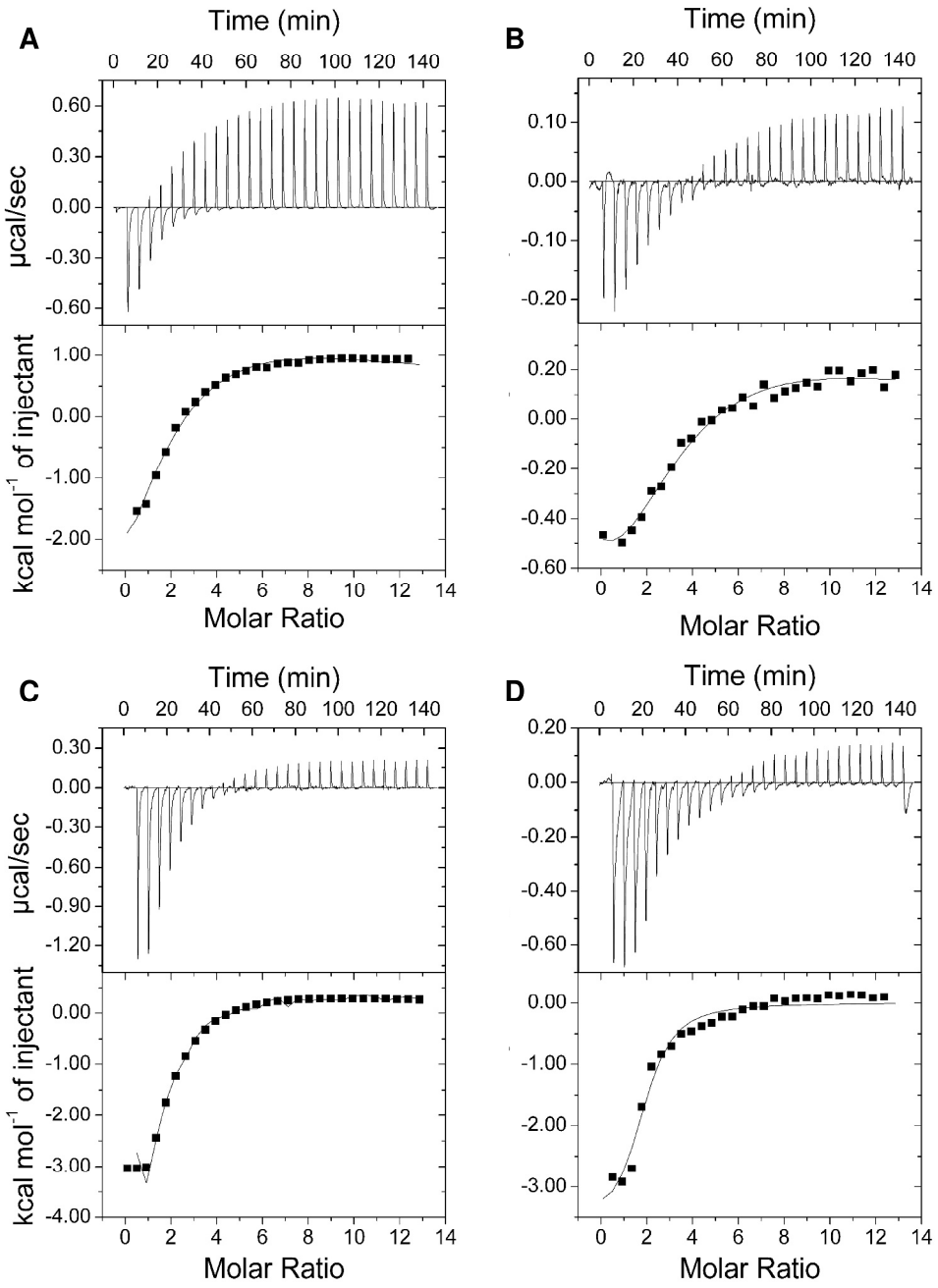


Fig. 1. ITC analysis of binding small effector molecules to PII (K58N) and (K58N/L56K) variants. The upper panels show the raw isothermal data in the form of the heat effect during the titration of PIIs solution (trimer concentration) with ligands. The lower panels show the binding isotherm and the best-fit curve according to the three sequential binding sites model. Titration of PII (A) (K58N) 33 µM with 2 mM ATP. (B) (K58N) 33 µM with 2 mM ADP. (C) (K58N/L56K) 33 µM with 2 mM ATP. (D) (K58N/L56K) 33 µM with 2 mM ATP.

Fig. 2. Enzymatic assay for SeNAGK in presence and absence of *Se*PII variants.

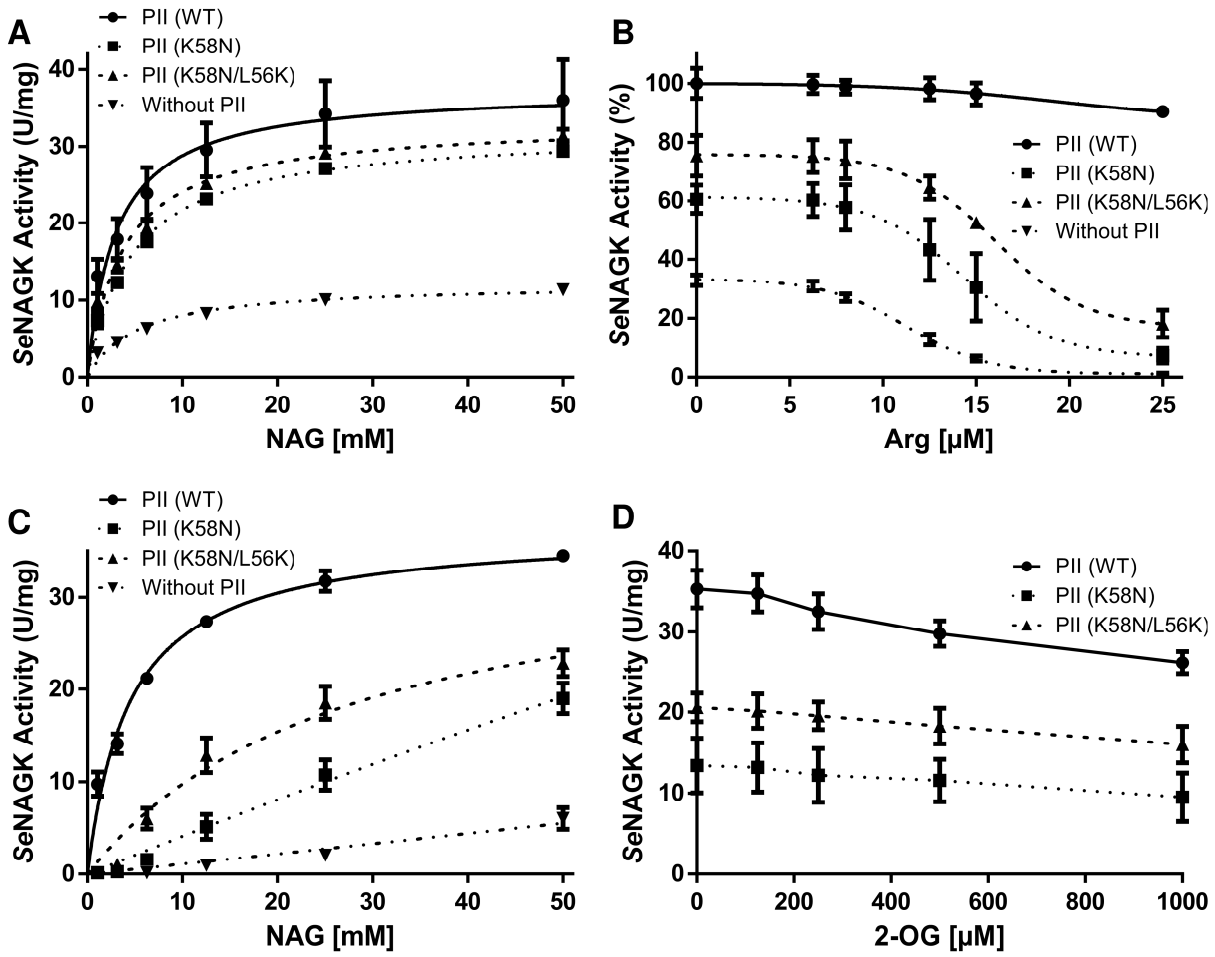


Fig. 2. NAGK Enzymatic assay with *Se*PII variants (K58N) and (K56N/L56K) in comparison to PII (WT). (A) The catalytic activity of *Se*NAGK with or without different variants of PII protein as indicated. NAG was used as a variable substrate. (B) Arginine-feedback inhibition of *Se*NAGK activity in the presence and absences of different variants of PII protein as indicated. Data were fitted according to a sigmoidal dose-response curve using a GraphPad Prism, yielding an IC_{50} for arginine. (C) The catalytic activity of *Se*NAGK with or without different variants of PII protein in presence of 11 μM Arg, as indicated. NAG was used as a variable substrate. (D) Effect of 2-OG on PII-based activation of NAGK in presence of 11 μM Arg, as indicated. SD as indicated by error bars, represents triplicate independent measurements.

Fig. 3. Complex formation between NAGK and *Se*PII (K58N) and (K58N/L56K) variants using SEC-MALS and SPR.

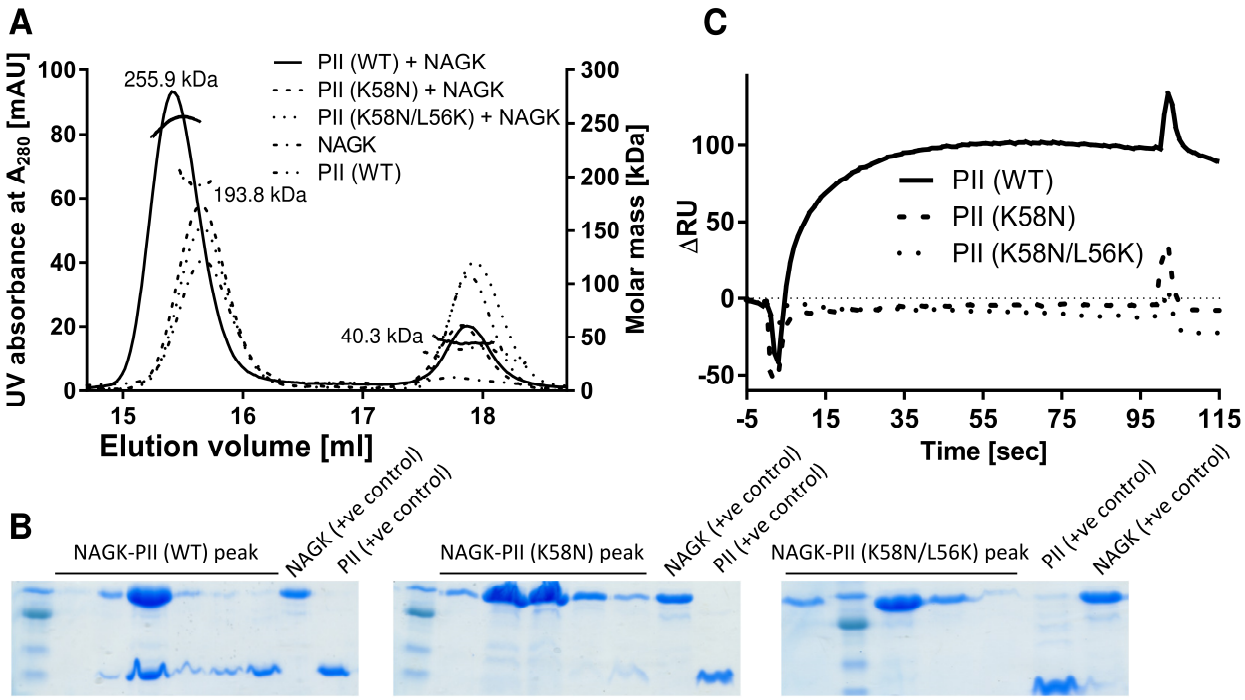


Fig. 3. PII-NAGK complex formation between *Se*NAGK and *Se*PII (K58N) and (K58N/L56K) variants using SEC-MALS and SPR. (A) SEC-MALS analysis shows dissociations of *Se*NAGK and *Se*PII (K58N) and (K58N/L56K) complexes through gel filtration run, while *Se*NAGK and *Se*PII (WT) form strong complex with higher molar mass. The eluted peaks from SEC runs were collected and subjected to SDS-PAGE. (B) SDS-PAGE analysis confirms presence of both PII (WT) and NAGK in the NAGK-PII (WT) complex peak, while NAGK was present alone for NAGK-PII (K58N) and NAGK-PII (K56N/L56K) peaks, indicating the dissociations of the complex during SEC for both PII variants. (C) SPR analysis for NAGK-PII complex formation. His_6-NAGK was immobilized to the Ni-NT sensor chip and different PIIs variants was injected to determine the resonance difference due to the PII-NAGK complex formation. PII (WT) was efficiently able to form complex with NAGK while PII variants were not able to complex with NAGK.

Fig. 4. SPR analysis for PipX-PII complex formation for *Se*PII (K58N) and (K58N/L56K) variants.

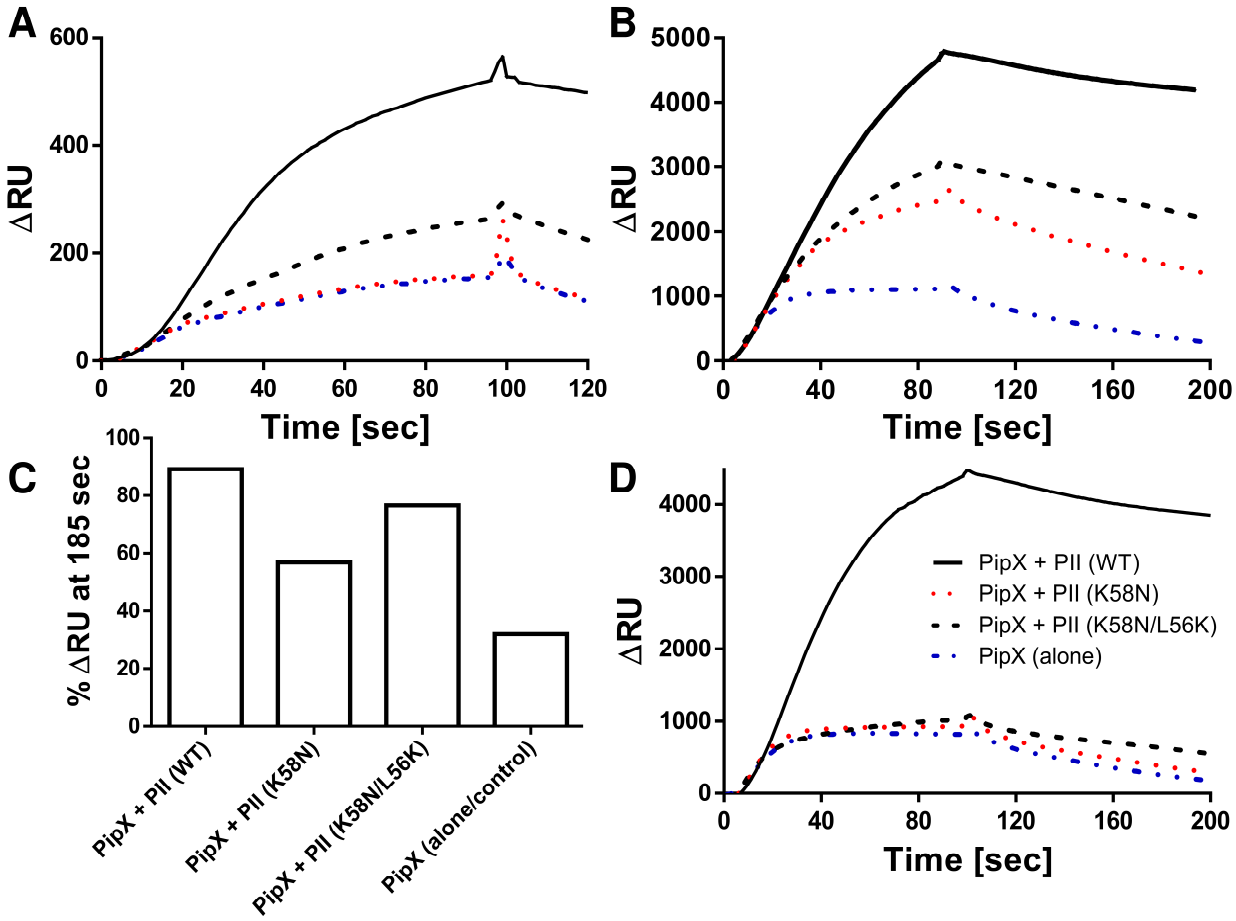


Fig 4. PII-PipX complex formation and dissociation for PII (K58N) and (K58N/L56K) variants in presents and absence of effector molecules. His_6-PipX (500nM) was injected to the chip in absence of PII and in presence of 500nM Strep-PIIs (WT), (K58N) and (K58N/L56K), (**A)** without effectors, **(B)** in presence of ADP (3mM), and **(D)** in presence of ATP (1mM). (**C)** shows the response signal in % at *t:185s* (90s after the end of the injection) of PipX in complex with different PII variants in presents 3 mM ADP. The response signal at the end of the injection at *t: 95* is normalized to 100%. The maintained signal at *t:185s* is an indicator for the stability of the complex.

Fig. 5. SPR of complex formation for PipX-PII (I86N) variant

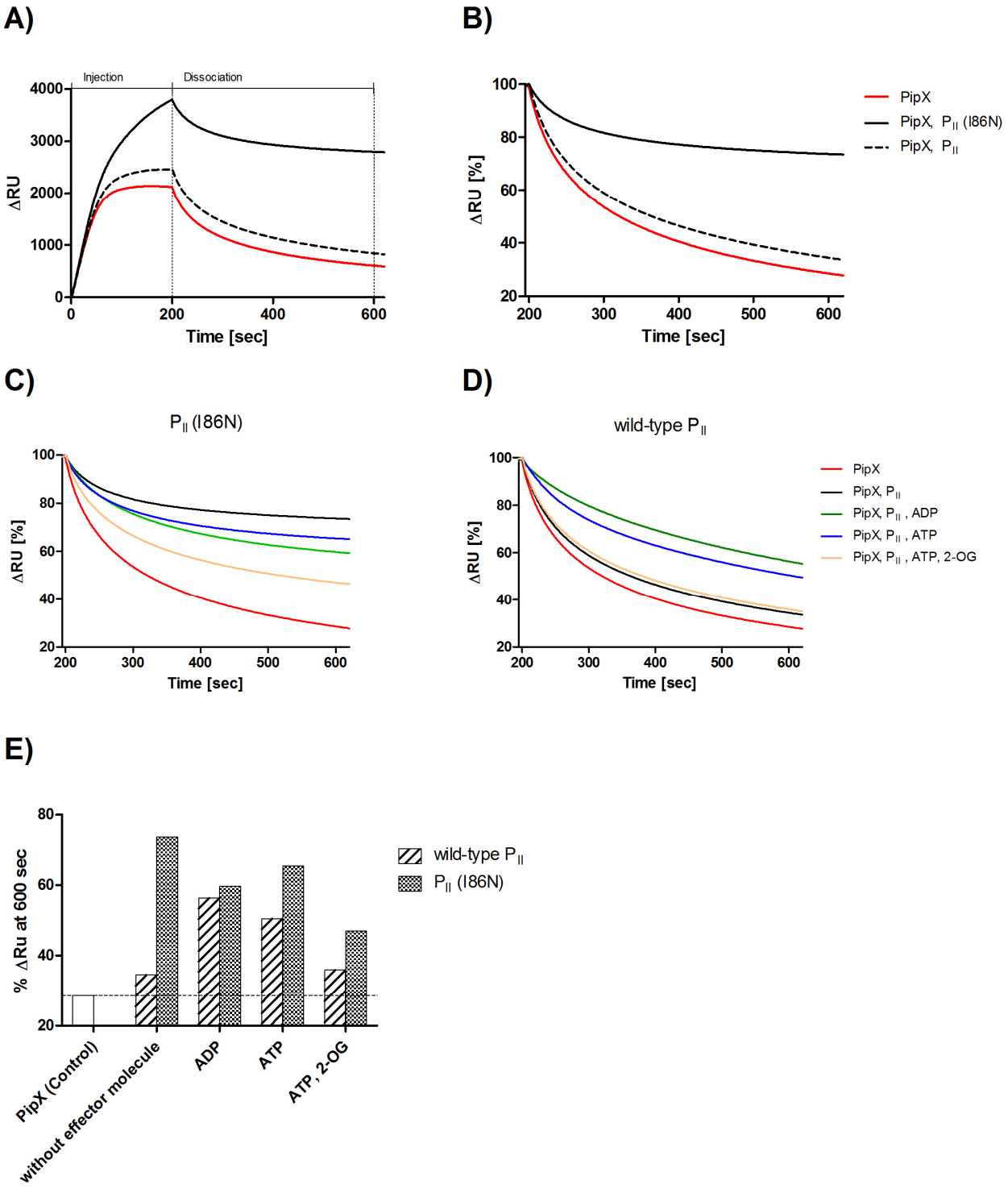
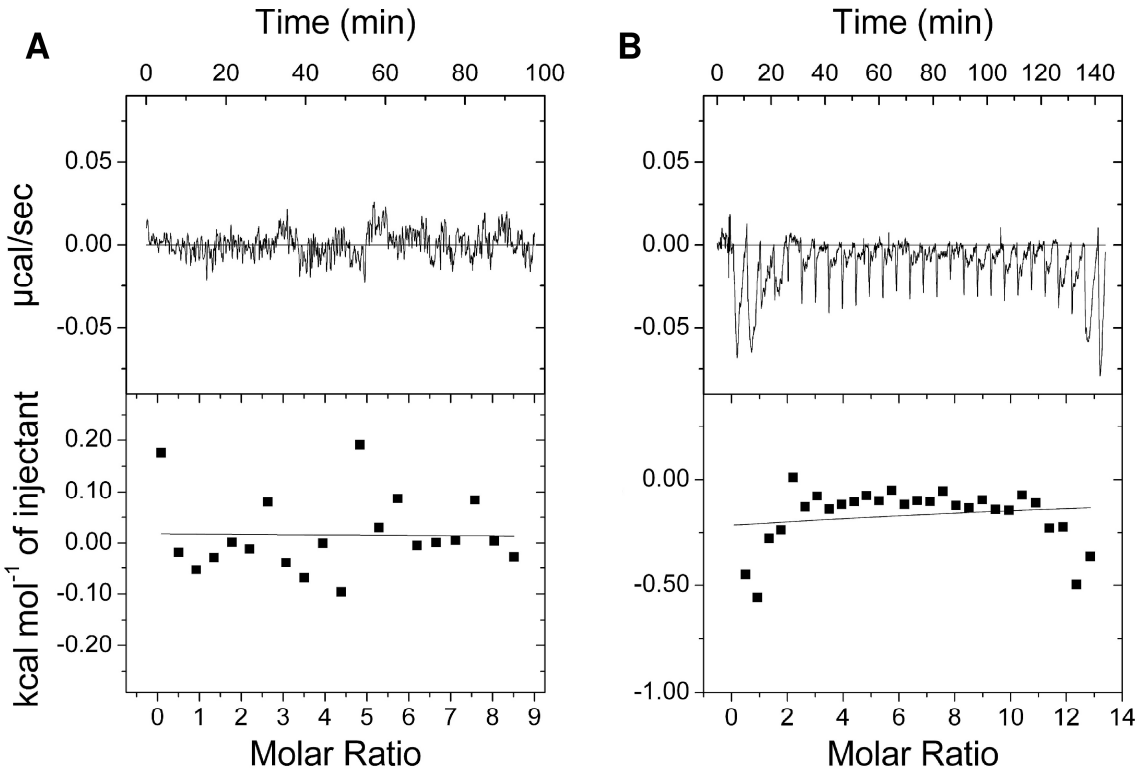


Fig 5. P_{II}-PipX complex formation and dissociation in presents and absence of effector molecules. **(A)** His_6-PipX (500nM) was injected to the chip in absence of P_{II} (red line) and in presence of 100nM Strep-P_{II} wild-type (black dashed line) and 100nM P_{II}(I86N) (black line), without effectors. Injection phase takes 200 s, followed by 400s dissociation. **(B)** Shows the dissociation of His_6-PipX (red line) alone and in

presents of Strep-P_{II} wild-type (black dashed line) and P_{II}(I86N) (black line). The response signal at the end of the injection at *t: 200s* is normalized to 100%. **(C)** and **(D)** shows the dissociation of P_{II}-PipX complex in presents of effectors, there for response signal at the end of the injection at *t: 200s* is normalized to 100%. In **C)** P_{II}(I86N) and in **D)** P_{II} wild-type was preincubated without effoctors (black line), 1mM ADP (green line), 1mM ATP (blue line) and 1mM ATP/ 1mM 2-OG (orange line). PipX without P_{II}, was used as control (red line). **(E)** shows the response signal in % at *t:600s* (400s after the end of the injection) of PipX in complex with P_{II} wild-type and P_{II}(I86N) in presents and absents of effectors. The maintained signal at *t:600s* is an indicator for the stability of the complex.

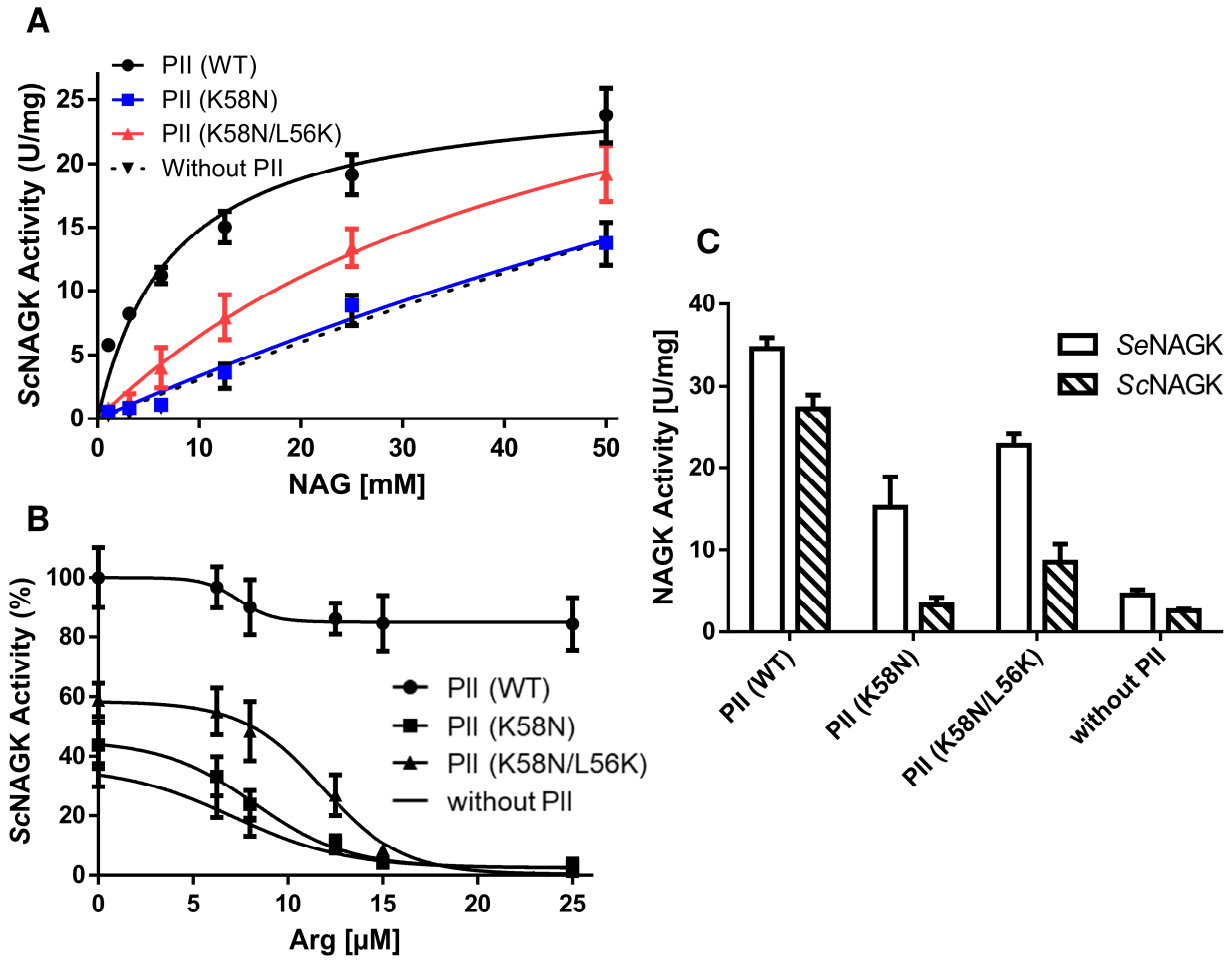
Supplementary Figures

Supplementary Fig. S1.



Supplementary Fig. S1. ITC analysis for binding of 2-OG (2mM) to PII (K58N) and (K58N/L56K) variants in presence of 1mM Mg-ATP. The upper panels show the raw isothermal data in the form of the heat effect during the titration of PIIs solution (trimer concentration) with ligands. Titration of PII (**A**) (K58N) 33 µM. (**B**) (K58N/L56K) 33 µM.

Supplementary Fig. S2. Enzymatic assay for determination of NAGKs activities in presence and absence of different PII variants.



Supplementary Fig. S2. NAGK Enzymatic assay with *Se*PII variants (K58N) and (K56N/L56K) in comparison to PII (WT). **(A)** The catalytic activity of *Sc*NAGK with or without different variants of PII protein, as indicated. NAG was used as a variable substrate. **(B)** Arginine-feedback inhibition of *Sc*NAGK activity in the presence and absences of different variants of PII protein, as indicated. Data were fitted according to a sigmoidal dose-response curve using a GraphPad Prism, yielding an IC_{50} for arginine. **(C)** The catalytic activity of different NAGKs with or without different variants of PII protein in presence of 12.5 µM of Arg, as indicated. SD as indicated by error bars, represents triplicate independent measurements.

Tables

Table 1: Dissociation constants (*Kd*) for binding ATP, ADP and 2-OG to the recombinant PII variants. The raw isothermal data were fitted according to three sequential binding sites models for PII trimer. NB: No binding.

Titrant/Protein	**Three-sites binding model**		
	K_{d1} **(µM)**	K_{d2} **(µM)**	K_{d3} **(µM)**
ATP			
PII (WT)*	(7.5)	(15.9)	(85.4)
PII (K58N)	27.3	90.2	1445.1
PII (K58N/L56K)	0.5	43.3	1016.26
ADP			
PII (WT)*	(10.6)	(19.3)	(133.4)
PII (K58N)	86.2	109.2	680.3
PII (K58N/L56K)	96.2	42.9	719.4
2-OG			
PII (WT)*	(2.9)	(6.4)	(83.0)
PII (K58N)	NB		
PII (K58N/L56K)	NB		

* Fokina O, Chellamuthu VR, Zeth K, Forchhammer K. A novel signal transduction protein P(II) variant from Synechococcus elongatus PCC 7942 indicates a two-step process for NAGK-P(II) complex formation. J Mol Biol. 2010b; 399(3):410-21.

Table 2: Kinetics of *Se*NAGK in presence and absence of PII variants. The kinetics of *Se*NAGK were determined in absence and presence of IC_{50} of Arg. NF: not fitted.

	K_m (mM) for NAG	V_{max}
In absence of feedback inhibitor Arg		
*Se*NAGK	5.13	12.12
*Se*NAGK + *Se*PII (WT)	2.99	37.46
*Se*NAGK + *Se*PII (K58N)	4.78	32.1
*Se*NAGK + *Se*PII (K58N/L56K)	3.77	33.19
In presence of feedback inhibitor Arg (11 µM)		
*Se*NAGK	NF	
*Se*NAGK + *Se*PII (WT)	4.48	37.28
*Se*NAGK + *Se*PII (K58N)	NF	
*Se*NAGK + *Se*PII (K58N/L56K)	27.78	36.74

2. Submitted manuscript

2.1. Manuscript 1 (Research Article):

Selim KA, Alva V, Hartmann MD, Forchhammer K. Structural and functional characterization of cyanobacterial PII-like protein CutA does not hint at an involvement in heavy metal tolerance. FEBS J. (Submitted 2019).

The FEBS Journal

Structural and functional characterization of cyanobacterial PII-like protein CutA does not hint at an involvement in heavy metal tolerance

Journal:	*The FEBS Journal*
Manuscript ID	Draft
Manuscript Type:	Regular Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Selim, Khaled; University of Tübingen Alva, Vikram; Max-Planck-Institute for Developmental Biology, Department of Protein Evolution Hartmann, Marcus; Max-Planck-Institute for Developmental Biology, Protein Evolution Forchhammer, Karl; University of Tuebingen, Inst. fur Mikrobiologie
Key Words:	

Submitted Manuscript for consideration in FEBS J.

Manuscript 1

Structural and functional characterization of cyanobacterial PII-like protein CutA does not hint at an involvement in heavy metal tolerance

Authors: Khaled A. Selim[a,b,1], Vikram Alva[b], Marcus D. Hartmann[b], Karl Forchhammer[a]

Affiliations:

[a]Interfaculty Institute for Microbiology and Infection Medicine, Organismic Interactions Department, Tübingen University, Auf der Morgenstelle 28, 72076 Tübingen, Germany.

[b]Protein Evolution Department, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany.

[1]To whom correspondence may be addressed. E-mail: KAS: Khaled.selim@uni-tuebingen.de

Authors E-mails list:

K.A. Selim: Khaled.selim@uni-tuebingen.de

V. Alva: vikram.alva@tuebingen.mpg.de

M.D. Hartmann: marcus.hartmann@tuebingen.mpg.de

K. Forchhammer: karl.forchhammer@uni-tuebingen.de

Introduction

The PII signal-transduction superfamily is widely spread in all domains of live, representing one of the largest and most ancient families of signaling proteins in nature. Signaling proteins of the PII superfamily are characterized by their highly conserved trimeric structure, consisting of a triangular core of β-sheets from the ferredoxin-like fold of the three subunits. Despite the highly conserved structure, the amino acid sequence conservation is quite low (Forchhammer & Lüddecke 2016, Selim et al. 2018).

One of PII-like protein, bearing the architectural core principle of PII proteins with the lowest sequence identity to canonical PII proteins (Forchhammer & Lüddecke 2016), is the divalent ion tolerance proteins CutA. CutA proteins are universally distributed in all domains of the life, including bacteria, archaea, plants, protozoa, animals, and humans. The cellular function of CutA proteins is still unclear. However, the CutA protein was first discovered in *E. coli* in a gene locus involved in divalent metal tolerance. In *E. coli*, the *cutA* locus contains three genes, in two operons, encoding cytoplasmic small protein of ~13 kDa (CutA1) in one operon and two inner membrane proteins (CutA2-3) in the second operon. The mutation in the *cutA* locus revealed increased sensitivity of *E. coli* cells to divalent metal ions (copper, zinc, nickel, cobalt and cadmium) owing to increased heavy metals uptake. Therefore, the CutA1 gene product was speculated to confer heavy meatal tolerance or to be involved in divalent cation homeostasis (Fong et al., 1995, Arnesano et al. 2003).

In mammals, CutA protein is localized in the brain cells for processing, trafficking and anchoring the membrane-bound neurotransmitter enzyme acetylcholinesterase (AChE), besides it supposed to control the copper homeostasis and ligand transport to membranes (Perrier et al. 2000, Arnesano et al. 2003, Yang et al. 2008, Liang et al. 2009). In human cell line, overexpression of CutA boosted the Cu^{2+} cytotoxicity by promoting Cu^{2+}-induced apoptosis (Yang et al. 2008). Moreover, the expression of CutA in animal cell line diminished the AChE levels, however still, there is no evidence for direct portion-protein interaction between AChE and CutA (Liang et al. 2009). Furthermore, the mammalian CutA was found to control the β-cleavage of β-amyloid precursor protein (APP) by interacting with β-site-APP cleaving protein 1 (BACE1). BACE1 is a putative β-secretase, which is the responsible for development of brain neurotoxic peptide β-amyloid (βA), the main pathogenesis reason for development of Alzheimer disease (Zhao et al. 2012). The knockdown and the overexpression of CutA in human cell line leaded to enhancement and decrease of BACE1-mediated APP processing/βA secretion, respectively (Zhao et al. 2012).

The first crystal structures of CutA proteins were determined form prokaryotic bacterium *E. coli* (*Ec*CutA; PDB: 1NAQ) and eukaryotic *Rattus norvegicus* (*RnCutA*; PDB: 1OSC) (Arnesano et al. 2003). Structurally both proteins were trimeric as well as in solution, highlighting the evolutionary conservation of the trimeric architecture of CutA proteins like canonical PII proteins (Arnesano et al. 2003). Each submit of the trimer is bearing a ferredoxin-like fold of β1α1β2-β3α2β4 with small β-hairpin extend connecting β2 andβ3 strands, and an additional small C-terminal: β5-strand and α3-helix (Arnesano et al. 2003). However, the CutA1 proteins do not hold the standard CXXC motif known for binding Cu-ions, but in *Ec*CutA crystal, Hg^{2+} ion was found bond within one of the inter cleft of two-subunits (Arnesano et al. 2003). Based on the pervious observations, the researchers concluded that CutA protein potentially implicated in divalent-metals homeostasis in *E. coli* (Fong et al., 1995, Arnesano et al. 2003).

To gain further insights about the distribution of CutA proteins among all domains of life, we performed excessive bioinformatic analysis to cluster the different subgroups of CutA protein. Next, we revised the PII superfamily to reveal the evolutionary relation between CutA and the other members of PII superfamily, with reflection on the distribution of PII superfamily proteins across different domains of life. Since, the function of CutA protein in cyanobacteria remains elusive. We chose members of cyanobacterial CutA proteins to study them physiologically, biochemically and structurally.

Results

Cluster analysis of CutA proteins reveals widespread distribution among all domains of the life.

The CutA protein is evolutionary quite conserved, with prokaryotic and eukaryotic homologs mostly sharing > 35% pairwise sequence identity, therefore we chose the *E. coli* CutA protein as a representative for our bioinformatic analysis. We searched the PDB70 profile-HMM database using the state-of-the-art remote homology detection method HHpred (Zimmermann 2018). While the best matches, as expected, were to CutA proteins from other organisms, we also found matches to many PII proteins at HHpred probability values of > 90%, suggesting that CutA and PII proteins share a common ancestry. Consistent with our observation, they are also classified under the same homology level in the SCOPe (Fox et al. 2014) and ECOD (Cheng et al. 2014) protein classification databases.

To comprehensively explore the taxonomic distribution and evolutionary conservation of CutA proteins, we searched the nr90 database, a version of the non-redundant protein sequence database filtered to a maximum pairwise sequence identity of 90%, for CutA homologs using PSI-

BLAST and investigated them using cluster analysis. The search yielded a total of 5736 homologs, which we subsequently clustered in CLANS based on the statistical significance of their all-against-all pairwise sequence similarities, as measured by PSI-BLAST P-values. The results revealed that the proteins of CutA family were established in the last universal common ancestor and have remained largely conversed during the course of evolution. Since, CutA proteins exhibit high pairwise sequence identities, we chose a stringent P-value cut-off to achieve separation between the archaeal, bacterial, and eukaryotic sequences (Fig. 1). At our chosen cut-off for clustering (1e-32), while the archaeal (colored green in the map) and bacterial (light red) sequences formed many loosely connected distinct clusters, exhibiting significant diversity, the eukaryotic (violet) sequences formed a single cluster, underpinning their high sequence conservation (the human and *E. coli* CutA proteins, for instance, exhibit a pairwise sequence identity of 43%). The eukaryotic CutA proteins are exist in most Eukaryotic lineages and primarily found in alveolates, metazoans, and green plants, but largely absent in Fungi. Whereas, the prokaryotic ones are distributed across almost all lineages of Bacteria and Archaea, including the deep-branching bacterial classes Actinobacteria and Cyanobacteria as well all the Asgard group of Archaea, which is thought to represent the closest prokaryotic relatives of eukaryotes. The cyanobacterial sequences (red) are mostly organized into two groups, with the exception of some that are scattered across the map. While, one of the two cyanobacterial groups are located close to the eukaryotic cluster, the other is further remote.

Relation of CutA proteins to the PII superfamily.

Since, our analysis suggested that CutA and PII proteins share a common ancestry (Fig. 1), we wanted next to identify the relation between CutA and other members of PII superfamily. To comprehensively explore the relationships between different protein families that adopt structurally the ferredoxin-like fold of PII proteins, we gathered homologs of representative PII/PII-like proteins of known structure and analyzed them using cluster analysis. To pick representatives for the homology searches, we chose the ECOD database. PII-like proteins are classified into 14 separate families within ECOD. We chose one representative per family and searched in the non-redundant protein sequence database for their homologs. The obtained sequences were pooled together, filtered down to a maximum pairwise sequence identity of 60%, and clustered in CLANS based on the statistical significance of their all-against-all pairwise sequence similarities. In the obtained map (Fig. 2), the members of the well-defined PII-family formed a central cluster (the canonical PII cluster) tightly connected to a cluster that contains the recently characterized SbtB and CPII proteins (Wheatley et al 2016, Selim et al. 2018).

The PII cluster contains: 1) The canonical PII proteins encoded by GlnB/GlnK/GlnZ which are associated with the main regulatory targets of glutamine synthetase (GS), glutamine-hydrolyzing NAD^+ synthetase, or ammonium transport channel (Amt); 2) The PotN-encoded PII proteins, which are found as a part of the polyamine spermidine biosynthesis pathway; 3) The PII proteins of the NifI clade that are found in pairs within the nitrogenase gene cluster; and 4) The PII proteins found in genetic association with heavy metal efflux pump (HmeP) that were previously termed PII-NG in the bioinformatic analysis by Sant'Anna et al. (2009). These four subgroups of the PII cluster remain connected to each other even at stringent P-value cut-offs, indicating that they are closely related. This justifies their designation as members of the PII family. The neighboring cluster, which contains SbtB, separates into three distinct subclusters at stringent P-value cut-offs (1e-29): The first subgroup contains the previously characterized SbtB proteins, which are found in association with the bicarbonate transporter SbtA (Shibata et al. 2002, Du et al. 2014, Selim et al. 2018); the second group comprises PII-like proteins that are associated with the carboxysome operon and which were termed CPII proteins (Wheatley et al 2016); the third subcluster comprises uncharacterized proteins found in genetic neighborhood with SbtA-like transporters (Shibata et al. 2002; we refer to them as SbtB-like proteins). Interestingly, the *sbtAB*-like operon exists as paralogue along with the *sbtAB* operon in some cyanobacterial strains, e.g. *Nostoc* sp. PCC 7120 possesses both types of *sbtAB* operons (Shibata et al. 2002), but until now the function of *SbtAB*-like operon remains enigmatic.

Four additional clusters of PII-like proteins radiate from the two central clusters. Of these, the tightest connections to the central clusters are made by two prokaryotic families of uncharacterized proteins that are composed solely of the ferredoxin-like fold of PII proteins, the DUF190 and DUF3240 families. While the DUF190 cluster is connected tightly to the SbtB cluster, DUF3240 is linked to the PII cluster, implying that they could be involved in inorganic carbon and nitrogen metabolisms, respectively. Of the remaining two, one cluster comprises the secondary messenger c-di-AMP receptor proteins DarA and PstA. These proteins are only found in bacteria and are absent in eukaryotes and archaea, possibly owing to c-di-AMP being a purely bacterial signaling molecule (Müller et al. 2015, Forchhammer & Lüddecke 2016). A further cluster, comprising of proteins of the uncharacterized DUF2179 family, radiates from the DarA/PstA cluster. Members of this cluster are present in archaea and bacteria and possess an N-terminal transmembrane domain. Another cluster, connected directly to the central clusters but distinct from PstA/DarA, contains the multi-domain NIF3 proteins (NGG1p-interacting factor 3). These proteins consist of a central PII-like domain, flanked on its N- and C-terminal ends by two NIF3-

like domains (Saikatendu et al. 2006, Godsey et al. 2007). NIF3 is ubiquitously conserved in all kingdoms of life; however, its exact cellular functions are still obscure. In eukaryotes, NIF3 proteins seem to be involved in transcriptional regulation, where its proposed role is the prevention of Ngg1p translocation to the nucleus by the formation of a complex with it in the cytoplasm (Saikatendu et al. 2006, Godsey et al. 2007, Fujishiro et al. 2014). In prokaryotes, it is likely that Nif3 proteins could also be involved in transcriptional regulation (Godsey et al. 2007). In fact, in *E. coli*, the NIF3-encoding homologue gene is highly upregulated upon genotoxic stress caused by DNA damage (Saikatendu et al. 2006). Recently, in *Methanocaldococcus jannaschii*, Nif3 protein was found in a gene cluster involved in the biosynthesis of the iron-guanylylpyridinol (FeGP) cofactor of [Fe]-hydrogenase and NIF3 was proposed to be an iron chaperone that is involved in the biosynthesis of the FeGP cofactor (Fujishiro et al. 2014). In additional to the central clusters, the NIF3 cluster is also connected to the universally distributed cluster of CutA proteins.

At the P-value cutoff chosen for clustering (1e-10), five groups do not make any connections to the central PII and SbtB clusters and represent their most distant homologs. These groups include HisG (ATP-phosphoribosyltransferase), N-acetylmuramoyl-L-alanine amidase (AmiC), two loosely connected clusters of proteases (Rhomboid_N and NRho), and DUF2007. While HisG, which is the first enzyme in biosynthetic pathway of histidine, is found in all domains of life (Supplementary Fig. S1), the others are exclusive to bacteria. In HisG, a C-terminal PII-like domain acts as a sensory domain and organizes the trimerization of the enzyme (Chellamuthu et al. 2013). HisG enzymes are allosterically inhibited by histidine and AMP. In the inter-subunit clefts of PII domain like the classical PII proteins, HisG can bind the allosteric inhibitor to mediate the inhibition of ATP-phosphoribosyltransferase activity (Chellamuthu et al. 2013). The member distribution of the clusters across different domains of life are summarized in supplementary Fig. (S1).

To this end, to examine the possibility of the involvement of cyanobacterial CutA proteins in heavy metal tolerance using biochemical and structural approaches, we chose a homolog from the group located close to eukaryotic subcluster, from the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942 (*Se*CutA) encoded by the ORF *Synpcc7942_2261*. As well as, one from the sequences scattered across the map (Fig. 1), close to other archaeal and bacterial sequences, from the filamentous cyanobacterium *Nostoc* sp. PCC 7120 (*Ns*CutA) encoded by the ORF *alr7093.*

CutA structure shows the typical PII architecture.

The predicted full-length polypeptide of *Ns*CutA is encoded by ORF *alr7093* (http://genome.microbedb.jp/cyanobase/) and consists of 104 amino acids with a calculated molecular mass of 12,024 Da (UniProt - Q8YL42 (Q8YL42_NOSS1)). For biochemical characterization, a recombinant C-terminal Strep-tagged CutA protein was expressed in *E. coli* and affinity-purified followed by size exclusion chromatography. Analytical gel filtration showed that *Ns*CutA eluted corresponding to the calculated mass of the trimeric recombinant protein and perfectly co-eluted with the well-studied canonical trimeric PII protein from *Synechococcus elongatus* PCC 7942 (*Se*PII), indicating that the recombinant CutA protein was prepared in its native trimeric state (Supplementary Fig. S2).

Up to now, no crystal structure was reported for cyanobacterial CutA. Therefore, we wanted to determine the structural basis for cyanobacterial CutA from the filaments cyanobacteria *Nostoc* sp. PCC 7120 via X-ray crystallography (*Ns*CutA). To this end, we set up crystallization trials for *Ns*CutA, which was rewarded with well-diffracting crystals. We obtained different crystal forms for CutA protein: an orthorhombic one in space group $P2_12_12_1$, containing two identical trimers in the asymmetric unit (ASU), two monoclinic crystals in space groups P2 and C2, and a trigonal one in P321, diffracting to 1.8Å, 2.1 Å, 2.3 Å, and 2.2Å, respectively.

We first worked with the orthorhombic crystal form and were able to solve its structure by molecular replacement using CutA structure from *Thermus thermophilus* (*Tt*CutA; PDB: 1V6H) as a search model. In ASU cell of the crystal, two homotrimers are existing to form a hexameric assembly in a bottom-to-bottom face mode with extensive contacts to create a dimer of trimers (Supplementary Fig. S3). The structure of *Ns*CutA protein is similar to overall structural topology as other known CutA proteins (Arnesano et al. 2003). As expected, the 3D topology of the refined *Ns*CutA structure shows the typical ferredoxin-like core architecture similar to the canonical PII proteins, with the tight cylindrical homotrimeric assembly shared among the PII superfamily (Fig 3*A&B*). All three subunits are essentially in the same conformation, and the whole *Ns*CutA trimer superimposes with Cα-rmsd (root-mean-square deviation) values of 0.99Å and 0.96Å on the *Tt*CutA and *Ec*CutA trimers respectively (Fig. 4, and Supplementary Fig. S4). The entire *Ns*CutA trimer superimposes with Cα rmsd values of 1.06Å and 1.09Å on the canonical *Se*PII trimer (PDB: 2xul) and PII-like protein SbtB (PDB:5O3P) (Fig. 4, Supplementary Fig. S4), which confirms a high degree of homology between different members of PII superfamily.

Notable overall differences to the structure of canonical PII proteins are found at the T-loop and C-terminus (Fig. 4*A*). Each monomer of the ferredoxin-like fold of CutA consists of two consecutive interlocking βαβ motifs, which form four central antiparallel β-sheet strands

surrounded at the outer surface by two antiparallel α-helices, with an additional helix α3 at the C-terminal end. In canonical PII proteins, the two βαβ motifs are connected between β2 and β3 strands by a long flexible loop called the T-loop (Fig. 3*B* and 4*A*). The T-loop, in most canonical PII structures, is disordered due to its high flexibility to adopt different conformations for signal-dependent target interaction (Forchhammer & Lüddecke 2016). However, in *Ns*CutA, as in other CutA homologues, the β2 and β3 strands are elongated compared to the corresponding strands in the canonical PII proteins and are linked only by a small hairpin turn of two-amino acid (Lys41, Gly42) (Fig. 3*B* and 4*A*). In *Ns*CutA, the β4 strand consists only of three amino acids (Val80-Ile82). Aside from that, there are two other loops emanating from both sides of the β4 strand: the B-loop, a small coil located between α2 helix and the start of β4 strand, and the C-loop, situated between the end of β4 strand and the C-terminal helix α3 (Fig. 3*B* and 4*A*). A remarkable difference between canonical PII proteins and CutA is found at the C-terminal end. The C-terminal helix α3 packs orthogonal to the N-terminus. Thus, the overall secondary structure arrangement of each monomer subunit of CutA is β1-α1-β2-β3-α2-(B-loop)-β4-(C-loop)-α3 (Fig. 3*B* and 4*A*), giving rise to a core architecture that perfectly matches the PII superfamily. The whole structural motif is invariant in all chains of the trimer and also formed in the structure of the other crystal forms that we solved subsequently.

The electrostatic potential surface of one face of *Ns*CutA is entirely negatively charged, whereas the other face has central negative cavity and the rest of that surface shows scattered negative, neutral and positive charges (Fig. 3*A*, upper panel). The electrostatic potential of interface of the lateral cleft between the trimer subunits is negatively charged (Fig. 3*A*, upper panel). In canonical PII proteins, the nucleotide binding pocket for ATP and ADP is located in the inter clefts between two subunits of the trimeric protein. Structural based alignment revealed that most of the residues that mediate the nucleotides binding in canonical PII proteins are not conserved in *Ns*CutA (Fig. 3*B* and 4*A*). However, in the *Ns*CutA trimer, the intersubunit cleft cavity is accessible via three solvent channels from outside. Therefore, the negatively charged clefts might mimic a functional property of metal channels or other positively charged cationic effectors.

Multiple sequence alignment of different members of CutA proteins (Supplementary Fig. S5) indicates that the intersubunit cleft cavity of CutA is formed by a number of conserved aromatic (Tyr38, Trp40, Tyr74, Tyr93, and Phe96) and charged residues (Glu47, Glu49, Glu75) (Fig. 4*C*), which suggests a conserved function for the inter clefts cavity of CutA proteins. Therefore, with the protein at hand, we wanted to define possible ligands able to bind CutA using isothermal titration calorimetry (ITC). *Ns*CutA protein was unable to bind to ATP, ADP, or unusual secondary massagers cAMP, which was not surprising as the binding pocket of CutA is too small to fit adenyl

nucleotides. Surprisingly, we were unable to detect also the binding of any metal ions including Cu^{2+} to CutA protein. Accordingly, we conclude that the possible binding of CutA to Cu^{2+} metal ions could be too weak to be detected with ITC. Therefore, we wanted to investigate the physiological cellular function of CutA protein under Cu^{2+} or other heavy metal stresses.

CutA doesn't mediate heavy metal tolerance in cyanobacteria.

To analyze the possible biological role of CutA in *Nostoc* sp. PCC 7120, we tried to create a single knockout mutant by deletion of ORF of *alr7093* gene. We were unable to obtain a complete segregated mutant, but instead we obtained only partially segregation, indicating that the *cutA* gene could be essential for *Nostoc* sp. PCC 7120 (Fig. 5). Alternatively, we decided to create a knockdown stain using antisense RNA strategy (Srivastava 2017, Bornikoel et al. 2018). In reverse-complement orientation, the *cutA* PCR amplicon (315 bp of *alr7093*) was cloned under the control of a strong copper inducible promoter (P_{petE}) into the promoter-less pAM1956 vector and transformed to *Nostoc* sp. PCC 7120 to obtain *alr7093* antisense knockdown strain (named *AS_NsCutA*). Downstream of P_{petE}-antisense cutA, a promoterless GFP reporter gene encoding for the green fluorescent protein (gfpmut2) is co-transcribed with antisense cutA (Fig. 5*A*). Genotypic analysis of *AS_NsCutA* strain confirmed that the strain contains P_{petE}-antisense cutA PCR fragment (Fig. 5*D*). Upon, Cu^{2+} addition to the media, the co-transcribed gfpmut2 gene resulted in emission of GFP fluorescence, which can be visualized by fluorescence microscopy indicating expression of *As_NsCutA* and GFP (Fig. 5*E*). Upon the transcription, the antisense-*cutA* RNA will hybridize with sense RNA of *cutA*, which should lead to knockdown of *alr7093* gene by preventing the translation of *cutA* gene product (Fig. 5*A*).

Additionally, we tried to knockout the homologues *cutA* gene in another model cyanobacterium organism, *Synechococcus elongatus* PCC 7942, by deletion of ORF *Synpcc7942_2261* (*ΔSeCutA*) to examine the possible physiological function of CutA in unicellular cyanobacteria (Fig. 5*B*). Surprisingly, unlike *Nostoc* sp. PCC 7120, genotypic analysis indicated that the *cutA* mutation is completely segregated, i.e. the wild-type gene fragment could not be detected by PCR (Fig. 5*F*). This finding implies that CutA is not essential for the viability of the unicellular cyanobacterium *S. elongatus* PCC 7942 under standard conditions, but it may be required for the filamentous life style of *Nostoc* sp. PCC 7120.

The predicted full length *Se*CutA polypeptide from *S. elongatus* PCC 7942, encoded by *Synpcc7942_2261* (Cyanobase) consists of 113 amino acids with a calculated molecular mass of 12,486 Da. The sequence identity between *Se*CutA and *Ns*CutA is quite low (~29%), suggesting that CutA protein could be evolutionary adapted to different life styles of cyanobacteria. To confirm

that the gene product of ORF *Synpcc7942_*2261 behaves as a trimer like *Ns*CutA protein, the C-terminal strep-tagged CutA protein of *S. elongatus* was purified and analyzed by analytical gel-filtration. The results revealed that *Se*CutA behaves as a trimer in solution identically to *Ns*CutA protein (Supplementary Fig. S2), implying that *Se*CutA is homologues to *Ns*CutA, despite of its low sequence identity.

To test whether CutA is involved in the tolerance to the heavy metals in cyanobacteria, cells of *S. elongatus* wild type*, Nostoc* sp. wild type (WT) and the mutant strains of *AS_NsCutA* and Δ*SeCutA* were tested against variety of divalent cations (Cu^{2+}, Fe^{2+}, Pb^{2+}, Co^{2+}, Zn^{2+}, Cr^{2+}, Cd^{2+}, Ni^{2+}, Mn^{2+}) in concentrations from 0-50 µM. Almost under all tested conditions in the presence of Cu^{2+}, the mutants behaved like the respective wild-type strain without notable difference. Only at higher concentrations of Cu^{2+} the *AS_NsCutA* strain showed a growth defect, but the mutant could recover from the metal stress as fast as the wild-type strain on normal BG11-agar plates (Fig. 6*A*). This indicates that this growth defect is likely due to accumulation of antisense *cutA* RNA, which is under the control of the copper inducible promoter P_{petE}, not due to the metal stress. In liquid culture, the *AS_NsCutA* mutant grew comparably to the WT in liquid cultures with/or without addition of 25 µM of Cu^{2+} (Fig. 6*B*), which further confirming the negligible effect of Cu^{2+} stress. The microscopic examination of Cu^{2+} treated and untreated WT and *AS_Ns*CutA mutant cells, did not reveal any notable difference (Fig. 6). The Δ*SeCutA* mutant showed only sensitivity against high concentration of Pb^{2+} (50 µM) (Fig. 6*C*). In nature, Pb^{2+} is normally one of the trace heavy metals and exist in very low concentrations, reveling that the observed phenotype of Δ*SeCutA* mutant cannot be environmentally reasonable. As a conclusion, we were unable to link the cyanobacterial CutA to the heavy metal tolerance. We found no difference between the cutA mutants and wild-type strains growing on BG11 media with either no heavy metal addition or with different concentrations of metals, indicating that CutA could play a different role in cyanobacteria rather than participating in heavy metals resistance.

Discussion

Evolutionary conservation of trimeric assembly of PII superfamily proteins.

PII-like proteins are structurally similar and clearly related to canonical PII proteins, but lacking PII PROSITE signature sequences and their functions were completely unknown until recently (Sant'Anna et al 2009, Forchhammer & Lüddecke 2016, Wheatley et al. 2016, Selim et al. 2018). Bioinformatics and structural genomics approaches expanded the members of the PII superfamily and proposed that the PII-like proteins represent an even more widespread family of trimeric regulators with ferredoxin-like fold, distributed in almost all living organisms (Sant'Anna et al 2009,

Forchhammer & Lüddecke 2016). Recently, new members of the PII superfamily have been characterized, suggesting that they are involved in controlling carbon metabolism. For example, the carboxysome-related PII protein, which binds ADP/AMP and bicarbonate and was proposed to sense bicarbonate availability (Wheatley et al. 2016). Recently, we revealed that a new PII-like protein in cyanobacteria, termed SbtB, senses the secondary messenger cAMP in addition to ATP, ADP and AMP. SbtB is involved in the control of the bicarbonate transporter SbtA and thereby links cAMP sensing to CO_2 metabolism (Selim et al. 2018). Furthermore, in firmicutes such as *Staphylococcus aureus* or *Bacillus subtillus*, a PII-like protein (termed PstA or DarA) was identified to sense the secondary messenger c-di-AMP; however, the exact function of that protein is still unknown (Gundlach et al. 2015, Campeotto et al. 2015**,** Müller et al. 2015). In all these cases, the PII/PII-like proteins exert their regulator function through binding small effector molecules, mainly adenylyl nucleotides, which induces conformational changes to transduce the signal to the PII/PII-like targets. In agreement, the trimeric packing of CutA proteins (Fig. 3 and Supplementary Fig. S4) with conserved lateral clefts, functionally unidentified to which small signaling molecules they response, resembles the ferredoxin-like arrangement of the well-studied signal transducer PII proteins. Given the low sequence conservation of PII-like proteins, even though, the superimposes of distantly unrelated proteins of PII superfamily matches perfectly the trimeric assembly of canonical PII protein with less than 1.0Å rmsd, highlighting evolutionary hallmark features of the trimeric architecture for the PII superfamily.

Addressing the evolutionary question of PII superfamily proteins is challenging due to diverse and small (around 110 amino acids) PIIs polypeptides sequences. Even though, it is strictly obvious that the proteins of PII superfamily share the same ancestor as the architectural principle of PII superfamily seems to be apparently widely distributed, while they evolutionary adapt to different sensory function. The conservation of trimeric assembly among the members of PII superfamily implies that they evolved probably early in history of the life independently from a trimeric ancestor protein. Then, they adapted later different sensory functions via differentiated ability to bind diverse of small effector molecules, to modulate different cellular functions through transducing the signals to the target interacting partners. Possibly as well, with increasing evolution pressure, some of PII-like protein arose from classical PII proteins via gene duplication events to adapt rapidly to new functions.

To our surprise, we were unable to show the ability of cyanobacterial CutA proteins neither to bind Cu^{2+} *in vitro* nor to confer heavy metal tolerance *in vivo*. The *cutA* mutations in two

independent cyanobacterial strains, *Nostoc* sp. and *S. elongatus*, revealed that CutA gene could be involved in different sensory function rather than metal sensing. Constant with our phenotypic observations, the level of gene expression of *cutA1* in plant pathogen *Xylella fastidiosa* showed no significance change in two different media under elevated levels of external Cu^{2+} concentrations (Rodrigues et al. 2008). Interestingly, also in phototrophic *Arabidiosis thaliana*, the *At*CutA knockout lines revealed that CutA is not essential for Cu^{2+} tolerance (Burkhead et al. 2003). Metal analysis of *Ec*CutA confirmed previously that the protein was in apo-state free of any tested dimetal ions (Cu, Ni, Zn, Mn, and Cd) (Arnesano et al. 2003). However, it's still possible that CutA plays a signaling or regulatory role related to heavy metals through the regulation of heavy metal efflux pumps similar to the other members of PII superfamily; like *GlnK* and *SbtB* which regulates ammonium or bicarbonate transport channels, respectively (Radchenko et al. 2010, Huergo et al. 2013). Thus, the CutA1 function may occur by ion binding (not likely) and/or by affecting ion import/export through interaction with membrane transporters.

So far, the creation of a completely segregated PII knockout mutant by SacB-based method in filamentous diazotrophic cyanobacteria were not successful, unless an additional copy of *glnB*-encoding for PII was provided in trans (Hanson et al. 1998, Zhang et al. 2007). However, in unicellular cyanobacteria, PII seems to be not essential, and can be knocked out, in particular, in a PipX-deficient background (Espinosa et al. 2009). Interesting that the same seems to be true for CutA and SbtB (encoded by OFE *all2133*; unpublished data) in *Nostoc* sp. PCC 7120, underlying the importance of PII-like proteins for filamentous lifestyle.

Altogether, based on our and previous structural studies of various proteins of PII superfamily (summarized in: Forchhammer & Lüddecke 2016, Wheatley et al. 2016, Selim et al. 2018), including canonical PII, SbtB, CPII, PstA, and CutA proteins, it seems that the trimeric assembly is a key evolutionary feature among PII superfamily. The conservation of trimeric structure, even when PII-domain is inserted in the protein with other domains like in case of NIF-3 and HisG clusters, makes the members of PII-superfamily a great toy for the use in protein design and for development of *in vivo* and *in vitro* biosensors. Remarkably, canonical PII proteins were used in several studies to develop 2-OG and ATP/ADP biosensors (Lüddecke & Forchhammer 2013, Lüddecke et al. 2017, Chen et al. 2018). Recently, CutA protein was used in protein engineering to design a symmetric tetrahedral protein assembly for co-assembling multi-component protein nanomaterials, which can be used in the drug delivery (King et al. 2014, Bale et al. 2015). Further structural, biochemical and physiological studies for understanding the sensory properties of cell

signaling PII superfamily, especially the DUF domains, are therefore important to understand the central metabolism, which could help for future biotechnological applications.

Methods

The bioinformatic analysis: The bioinformatic analysis was carried out using tools in the MPI-Bioinformatics Toolkit (Zimmermann et al. 2018). HHpred search was performed seeded with the CutA protein from *E. coli* against the PDB_mmCIF70 database, a version of the PDB database clustered to a maximum pairwise sequence identity of 70%. To gather CutA homologs, we searched the nr90 sequence databases using two iterations of PSI-BLAST, with an E-value inclusion cut-off of 1E-3. To pick representatives for PII homology, we chose the ECOD database. The obtained full-length sequences were clustered in CLANS at a P-value cut-off of 1e-32, with attract value = 2 and repulse value = 20. To generate PII cluster, the sequences were filtered down to a maximum pairwise sequence identity of 60%.

Isothermal titration calorimetry (ITC): ITC was performed using VP-ITC (MicroCal) as described previously in (Selim et al. 2018), after excessive dialysis of CutA protein in Tris-HCl buffer (pH 7.8) supplemented with 200 mM NaCl and Chelex 100 to remove any contamination of divalent cations.

Size exclusion chromatography (SEC): Analytical SEC was performed using ÄKTA chromatography system at room temp on the Superdex™ 200 column PC 3.2/30 (GE Healthcare, with geometric volume 2.4 ml), as described previously in (Selim et al. 2018).

Crystallization: The crystallization of CutA protein was achieved via vapor diffusion at 20 ℃ in 96-well sitting-drop plates, as described previously in (Selim et al. 2018). The CutA structure of orthorhombic crystal was solved by molecular replacement using MOLREP and the *Thermus thermophilus* CutA structure (PDB: 1V6H) as a search model. The CutA structure was completed by cyclic manual modeling using Coot program and refinement via REFMAC5 program. The structure requires further refinements; therefore, the statistics is not included.

Molecular biology methods: The C-terminal recombinant Strep-tagged CutA protein from *Nostoc* sp. PCC7120 (encoded by ORF *alr7093*) was cloned into pASK-IBA3 vector, as described previously in (Selim et al. 2018), using primers 1414 and 1415. The expression and purification of recombinant CutA protein was achieved, as described previously in (Selim et al. 2018). For creation of knockout mutant in the ORF *Synpcc7942_2261* encoding for CutA in *Synechococcus elongatus* PCC 7942, kanamycin resistance cassette (using: 1238 & 1239 primers) and upstream (using: 1240 & 1241 primers) and downstream (using: 1242 & 1243 primers) regions of *Synpcc7942_2261* were amplified and cloned into PUC19 plasmid, the resulted plasmid encoding for CutA-knockout construct was transformed into *Synechococcus elongatus* PCC 7942, as described previously for *Synechocystis* sp. PCC 6803 in (Selim et al. 2018). The mutants were selected on BG11-plates supplemented with 50 µg/ml kanamycin. For creation of knockdown mutant

in *Nostoc* sp. PCC7120, *alr7093* encoding for CutA was cloned in reverse-complement orientation (using: 1178 & 1179 primers) into pAM1956 vector under the control of a strong copper inducible promoter (P_{petE}) (using: Fw & 1177 primers), as described in (Srivastava 2017, Bornikoel et al. 2018). The primers used in this study are listed in Table (S1).

References

1. Arnesano F, Banci L, Benvenuti M, Bertini I, Calderone V, Mangani S, Viezzoli MS. The evolutionarily conserved trimeric structure of CutA1 proteins suggests a role in signal transduction. J Biol Chem. 2003; 278(46):45999-6006.
2. Bale JB, Park RU, Liu Y, Gonen S, Gonen T, Cascio D, King NP, Yeates TO, Baker D. Structure of a designed tetrahedral protein assembly variant engineered to have improved soluble expression. Protein Sci. 2015; 24(10):1695-701.
3. Bornikoel J, Staiger J, Madlung J, Forchhammer K, Maldener I. LytM factor Alr3353 affects filament morphology and cell-cell communication in the multicellular cyanobacterium *Anabaena* sp. PCC 7120. Mol Microbiol. 2018; 108(2):187-203.
4. Burkhead JL, Abdel-Ghany SE, Morrill JM, Pilon-Smits EA, Pilon M. The *Arabidopsis thaliana* CUTA gene encodes an evolutionarily conserved copper binding chloroplast protein. Plant J. 2003; 34(6):856-67.
5. Campeotto I, Zhang Y, Mladenov MG, Freemont PS, Gründling A. Complex structure and biochemical characterization of the *Staphylococcus aureus* cyclic diadenylate monophosphate (c-di-AMP)-binding protein PstA, the founding member of a new signal transduction protein family. J Biol Chem. 2015; 290(5):2888-901.
6. Chellamuthu VR, Alva V, Forchhammer K. From cyanobacteria to plants: conservation of PII functions during plastid evolution. Planta. 2013; 237(2):451-62.
7. Chen HL, Latifi A, Zhang CC, Bernard CS. Biosensors-Based In Vivo Quantification of 2-Oxoglutarate in Cyanobacteria and Proteobacteria. Life (Basel). 2018; 8(4).
8. Cheng H, Schaeffer RD, Liao Y, Kinch LN, Pei J, Shi S, Kim BH, Grishin NV. ECOD: an evolutionary classification of protein domains. PLoS Comput Biol. 2014; 10(12):e1003926.
9. Du J, Förster B, Rourke L, Howitt SM, Price GD. Characterisation of cyanobacterial bicarbonate transporters in *E. coli* shows that SbtA homologs are functional in this heterologous expression system. PLoS One. 2014; 9(12):e115905.
10. Espinosa J, Castells MA, Laichoubi KB, Contreras A. Mutations at pipX suppress lethality of PII-deficient mutants of Synechococcus elongatus PCC 7942. J Bacteriol. 2009; 191(15):4863-9.
11. Fong ST, Camakaris J, Lee BT. Molecular genetics of a chromosomal locus involved in copper tolerance in *Escherichia coli* K-12. Mol Microbiol. 1995; 15(6):1127-37.
12. Forchhammer K, Lüddecke J. Sensory properties of the PII signalling protein family. FEBS J. 2016; 283(3):425-37.

13. Fox NK, Brenner SE, Chandonia JM. SCOPe: Structural Classification of Proteins--extended, integrating SCOP and ASTRAL data and classification of new structures. Nucleic Acids Res. 2014; 42:D304-9.
14. Fujishiro T, Ermler U, Shima S. A possible iron delivery function of the dinuclear iron center of HcgD in [Fe]-hydrogenase cofactor biosynthesis. FEBS Lett. 2014; 588(17):2789-93.
15. Godsey MH, Minasov G, Shuvalova L, Brunzelle JS, Vorontsov II, Collart FR, Anderson WF. The 2.2 A resolution crystal structure of *Bacillus cereus* Nif3-family protein YqfO reveals a conserved dimetal-binding motif and a regulatory domain. Protein Sci. 2007; 16(7):1285-93.
16. Gundlach J, Dickmanns A, Schröder-Tittmann K, Neumann P, Kaesler J, Kampf J, Herzberg C, Hammer E, Schwede F, Kaever V, Tittmann K, Stülke J, Ficner R. Identification, characterization, and structure analysis of the cyclic di-AMP-binding PII-like signal transduction protein DarA. J Biol Chem. 2015; 290(5):3069-80.
17. Hanson TE, Forchhammer K, Tandeau de Marsac N, Meeks JC. Characterization of the glnB gene product of *Nostoc punctiforme* strain ATCC 29133: glnB or the PII protein may be essential. Microbiology. 1998; 144:1537-47.
18. Huergo LF, Chandra G, Merrick M. P(II) signal transduction proteins: nitrogen regulation and beyond. FEMS Microbiol Rev. 2013; 37(2):251-83.
19. King NP, Bale JB, Sheffler W, McNamara DE, Gonen S, Gonen T, Yeates TO, Baker D. Accurate design of co-assembling multi-component protein nanomaterials. Nature. 2014; 510(7503):103-8.
20. Liang D, Nunes-Tavares N, Xie HQ, Carvalho S, Bon S, Massoulié J. Protein CutA undergoes an unusual transfer into the secretory pathway and affects the folding, oligomerization, and secretion of acetylcholinesterase. J Biol Chem. 2009; 284(8):5195-207.
21. Lüddecke J, Forchhammer K. From PII signaling to metabolite sensing: a novel 2-oxoglutarate sensor that details PII-NAGK complex formation. PLoS One. 2013; 8(12):e83181.
22. Lüddecke J, Francois L, Spät P, Watzer B, Chilczuk T, Poschet G, Hell R, Radlwimmer B, Forchhammer K. P_{II} Protein-Derived FRET Sensors for Quantification and Live-Cell Imaging of 2-Oxoglutarate. Sci Rep. 2017; 7(1):1437.
23. Müller M, Hopfner KP, Witte G. c-di-AMP recognition by *Staphylococcus aureus* PstA. FEBS Lett. 2015; 589(1):45-51.
24. Perrier AL, Cousin X, Boschetti N, Haas R, Chatel JM, Bon S, Roberts WL, Pickett SR, Massoulié J, Rosenberry TL, Krejci E. Two distinct proteins are associated with tetrameric acetylcholinesterase on the cell surface. J Biol Chem. 2000; 275(44):34260-5.
25. Radchenko MV, Thornton J, Merrick M. (2010) Control of AmtB-GlnK complex formation by intracellular levels of ATP, ADP, and 2-oxoglutarate. J Biol Chem. 2010 Oct 1;285(40):31037-45.
26. Rodrigues CM, Takita MA, Coletta-Filho HD, Olivato JC, Caserta R, Machado MA, de Souza AA. Copper resistance of biofilm cells of the plant pathogen *Xylella fastidiosa*. Appl Microbiol Biotechnol. 2008; 77(5):1145-57.

27. Saikatendu KS, Zhang X, Kinch L, Leybourne M, Grishin NV, Zhang H. Structure of a conserved hypothetical protein SA1388 from *S. aureus* reveals a capped hexameric toroid with two PII domain lids and a dinuclear metal center. BMC Struct Biol. 2006; 6:27.
28. Sant'Anna FH, Trentini DB, de Souto Weber S, Cecagno R, da Silva SC, Schrank IS. The PII superfamily revised: a novel group and evolutionary insights. J Mol Evol. 2009; 68(4):322-36.
29. Selim KA, Haase F, Hartmann MD, Hagemann M, Forchhammer K. P_{II}-like signaling protein SbtB links cAMP sensing with cyanobacterial inorganic carbon response. Proc Natl Acad Sci U S A. 2018; 115(21):E4861-E4869.
30. Shibata M, Katoh H, Sonoda M, Ohkawa H, Shimoyama M, Fukuzawa H, Kaplan A, Ogawa T. Genes essential to sodium-dependent bicarbonate transport in cyanobacteria: function and phylogenetic analysis. J Biol Chem. 2002; 277(21):18658-64.
31. Srivastava A, Brilisauer K, Rai AK, Ballal A, Forchhammer K, Tripathi AK. Down-Regulation of the Alternative Sigma Factor SigJ Confers a Photoprotective Phenotype to *Anabaena* PCC 7120. Plant Cell Physiol. 2017; 58(2):287-297.
32. Wheatley NM, Eden KD, Ngo J, Rosinski JS, Sawaya MR, Cascio D, Collazo M, Hoveida H, Hubbell WL, Yeates TO. A PII-Like protein regulated by bicarbonate: structural and biochemical studies of the carboxysome-associated CPII protein. J Mol Biol. 2016; 428(20):4013-4030.
33. Yang J, Li Q, Yang H, Yan L, Yang L, Yu L. Overexpression of human CUTA isoform2 enhances the cytotoxicity of copper to HeLa cells. Acta Biochim Pol. 2008; 55(2):411-5.
34. Zhang Y, Pu H, Wang Q, Cheng S, Zhao W, Zhang Y, Zhao J. PII is important in regulation of nitrogen metabolism but not required for heterocyst formation in the Cyanobacterium *Anabaena* sp. PCC 7120. J Biol Chem. 2007; 282(46):33641-8.
35. Zhao Y, Wang Y, Hu J, Zhang X, Zhang YW. CutA divalent cation tolerance homolog (*Escherichia coli*) (CUTA) regulates β-cleavage of β-amyloid precursor protein (APP) through interacting with β-site APP cleaving protein 1 (BACE1). J Biol Chem. 2012; 287(14):11141-50.
36. Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol. 2018; 430(15):2237-2243.

Figures

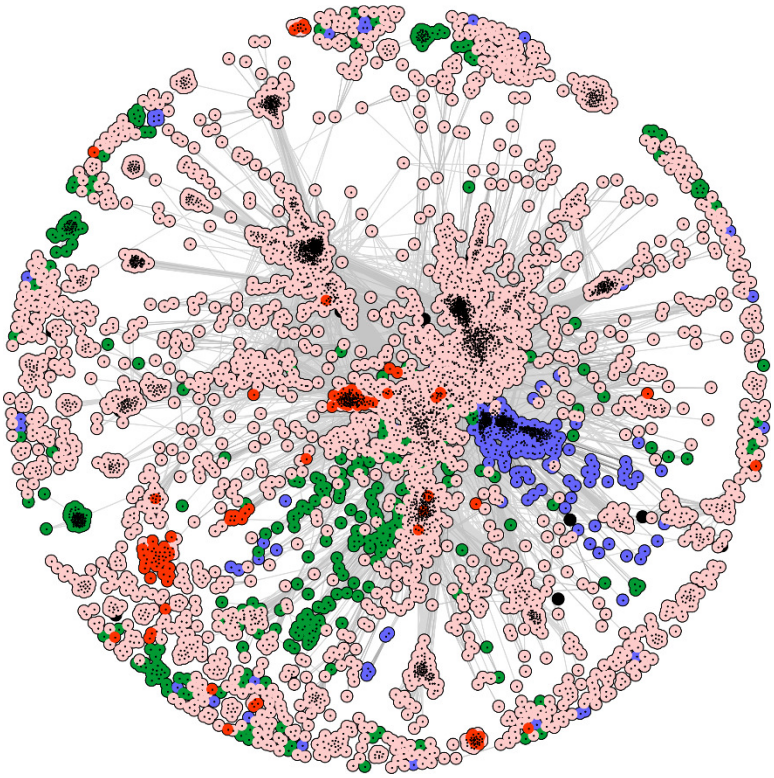


Fig. 1. Cluster mapping of CutA proteins revealed a widespread of CutA proteins among all domains of life. Bacterial sequences are colored in light red, cyanobacterial ones in red, archaeal sequences in green, and eukaryotic sequences in violet.

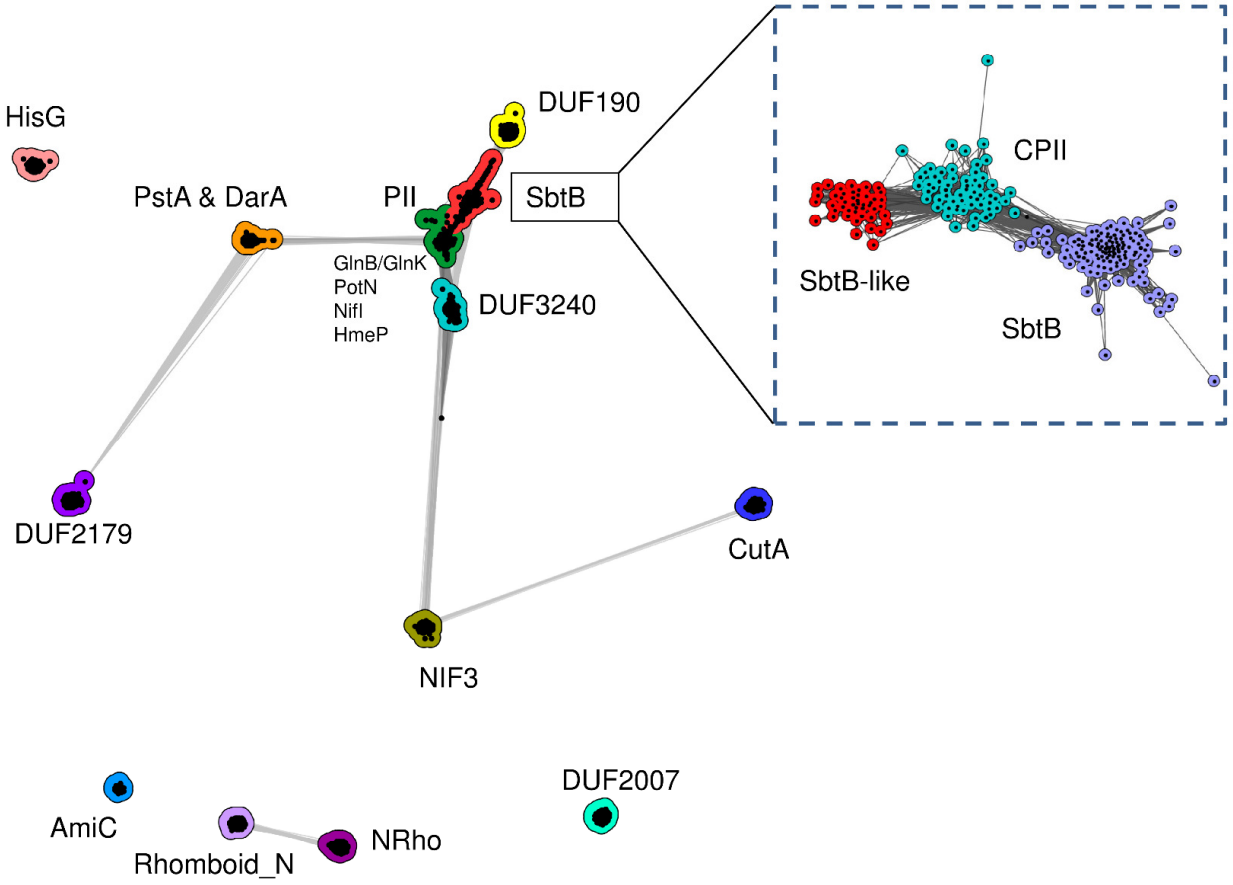


Fig. 2. Cluster mapping of PII superfamily. Cluster of canonical PII proteins is colored in green (including: glnB and glnK, PotN, NifI, and HmeP encoded PII proteins), cluster of SbtB proteins is colored in red; the inset highlight the subgroups of SbtB cluster (including: classical SbtB, SbtB-like and CPII proteins), cluster of c-di-AMP binding proteins PstA and DarA is colored in orange, Cluster of CutA proteins is colored in blue, cluster of NIF3 proteins is colored in green, and clusters including DUF domains 190, 3240, and 2179 are colored in yellow, light blue, and violet, respectively. The outliers clusters include HisG, AmiC, Rhomboid_N, NRho, and DUF2007.

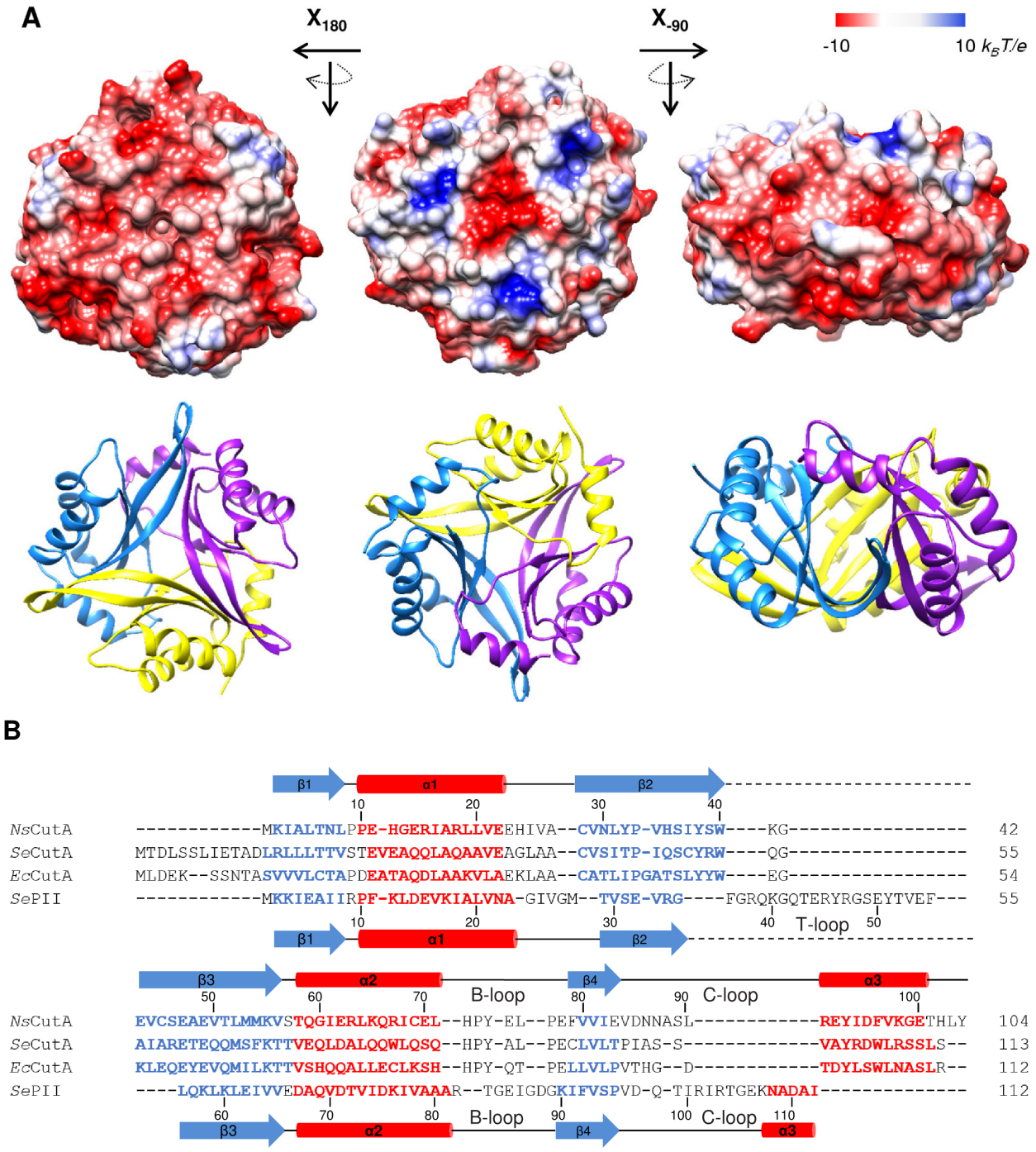


Fig. 3. Structural features of *Ns*CutA protein. (A) Overall structural architectural of *Ns*CutA. The structure of the trimeric *Ns*CutA is shown as an electrostatic potential surface (upper panel) and as a cartoon (lower panel). (B) Structural based alignment of canonical PII protein (*Se*PII) and CutA proteins (*Ns*CutA, *Se*CutA, and *Ec*CutA).

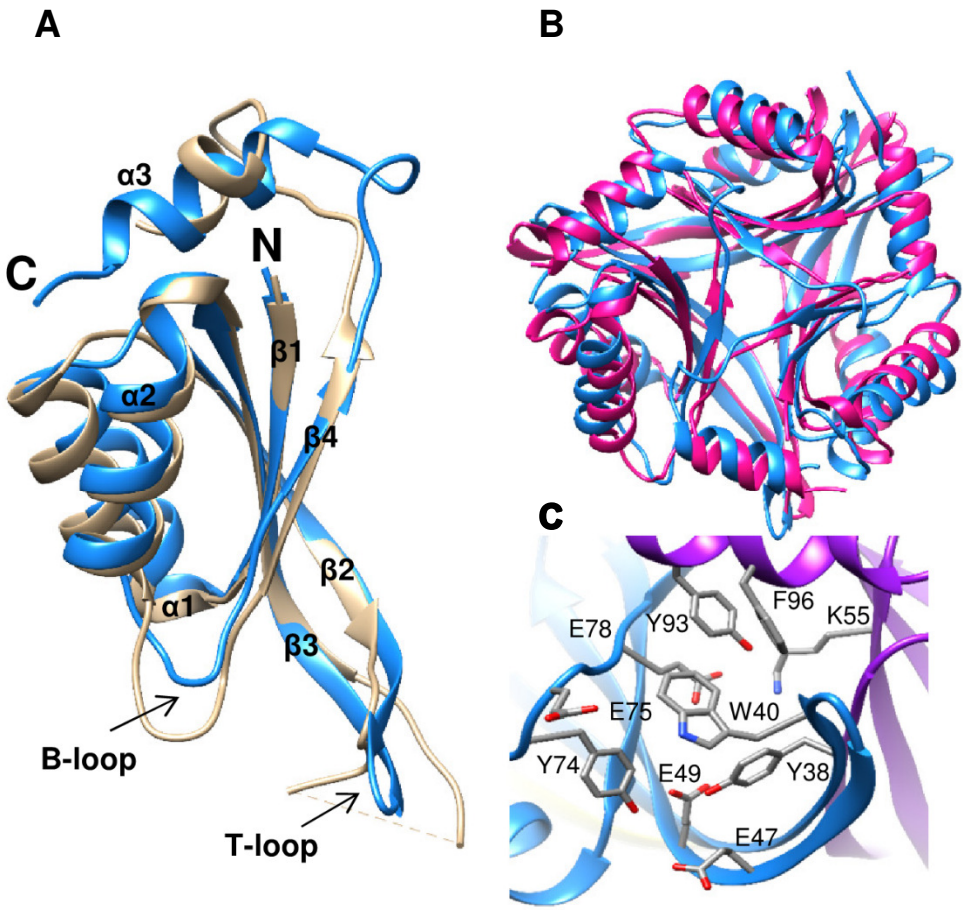


Fig. 4. Structural features of *Ns*CutA protein. (A) Superposition of monomeric subunits of CutA (blue) with *Se*PII (brown; PDB ID code 2XUL), yielding a 1.06-Å rmsd. The secondary structure elements and the characteristic structural motifs are indicated. (B) Superposition of trimeric *Ns*CutA (blue) over *Ec*CutA (pink; PDB ID code 1NAQ), yielding a 0.96 Å rmsd. (C) Close-up of the intersubunits cleft showing the conserved residues building up the CutA cavity.

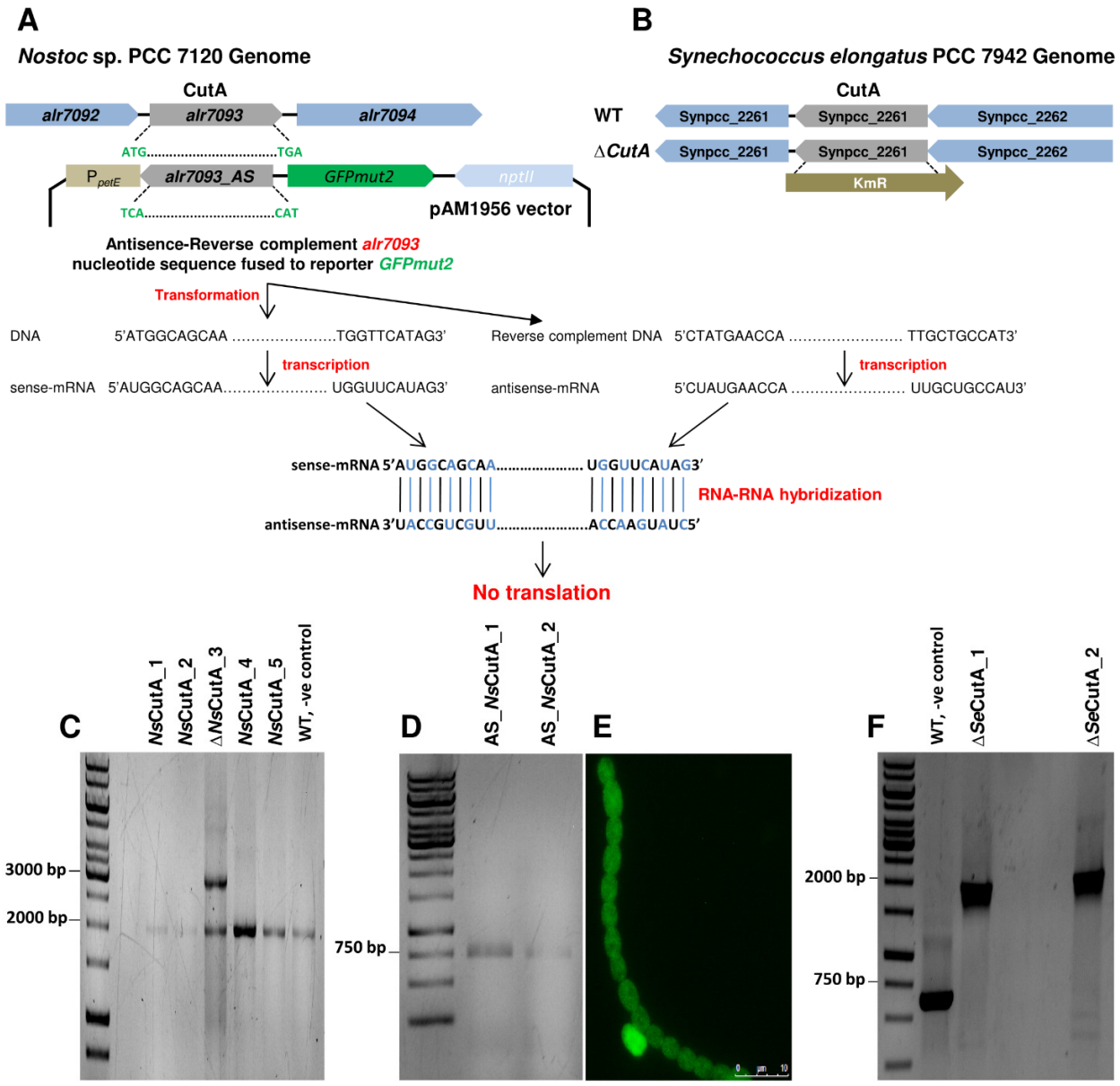


Fig. 5. Genotypic characterization of knockdown and knockout CutA mutants. (A) Schematic representation of genetic organization of *alr7093* (designated *cutA*) gene in the *Nostoc* sp. PCC7120 genome, the knockdown of the *cutA* gene was accomplished by cloning of ORF *alr7093* in reverse completed order under the control of PetE promoter. (B) Schematic representation of genetic organization of *Synpcc7942_2261* (designated *cutA*) gene in the *Synechococcus elongatus* PCC 7942 genome, the deletions of the *cutA* gene was accomplished by replacement of ORF *Synpcc7942_2261* with kanamycin resistance cassette. (C & F) PCR showing an incomplete (C) and a complete (F) segregation of knamycin-resistance-cassette within Δ*cutA* knockouts within *Nostoc* sp. and *S. elongatus*, respectively. (D) PCR shows presence of cutA-*alr7093* gene under the control of PetE-promoter within the AS_*Ns*CutA strains. (E) Microscopic examination of antisense knockdown strain reveals expression of GFP within the AS_*Ns*CutA strain; indicating the expression of antisense-cutA.

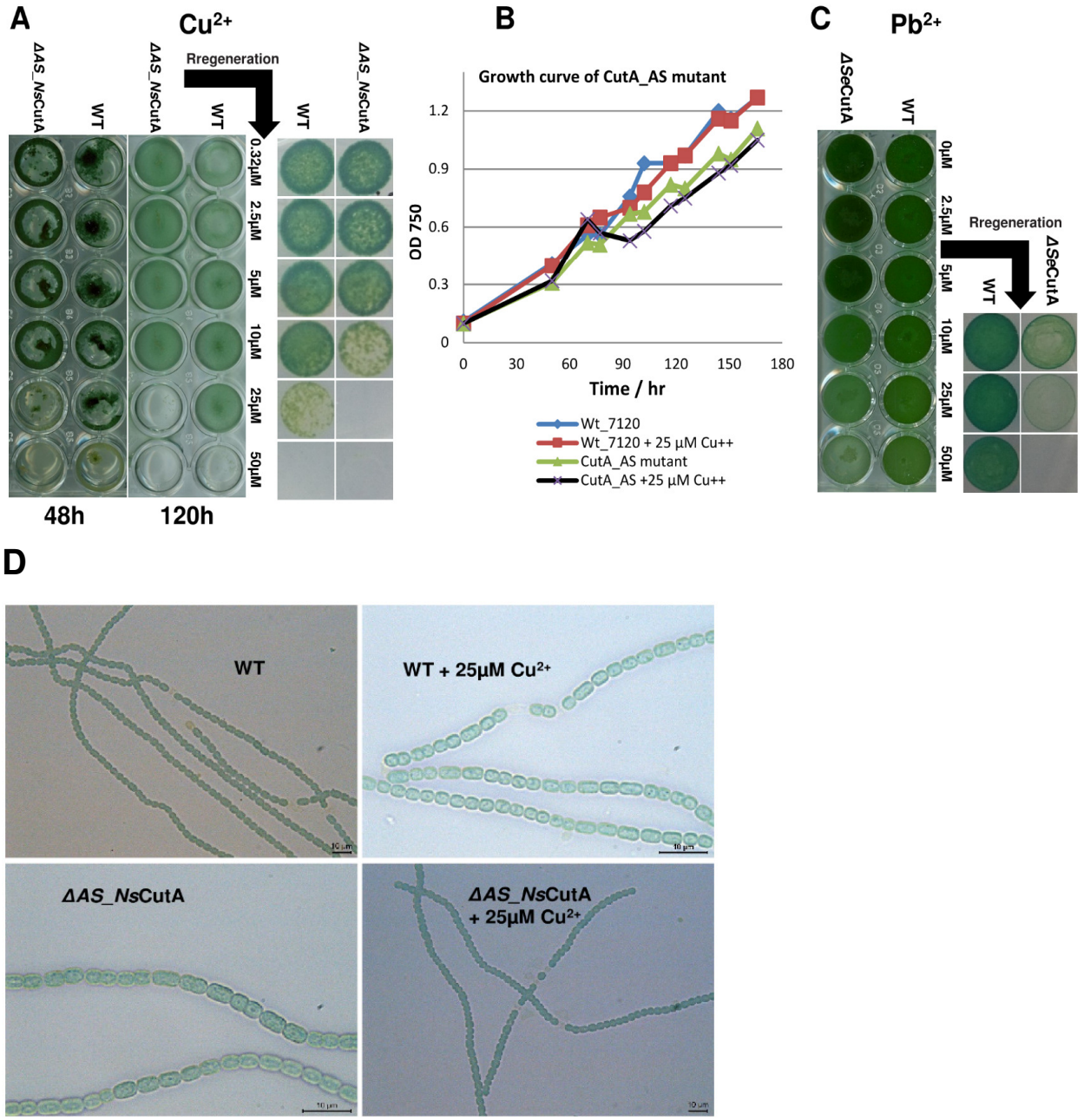
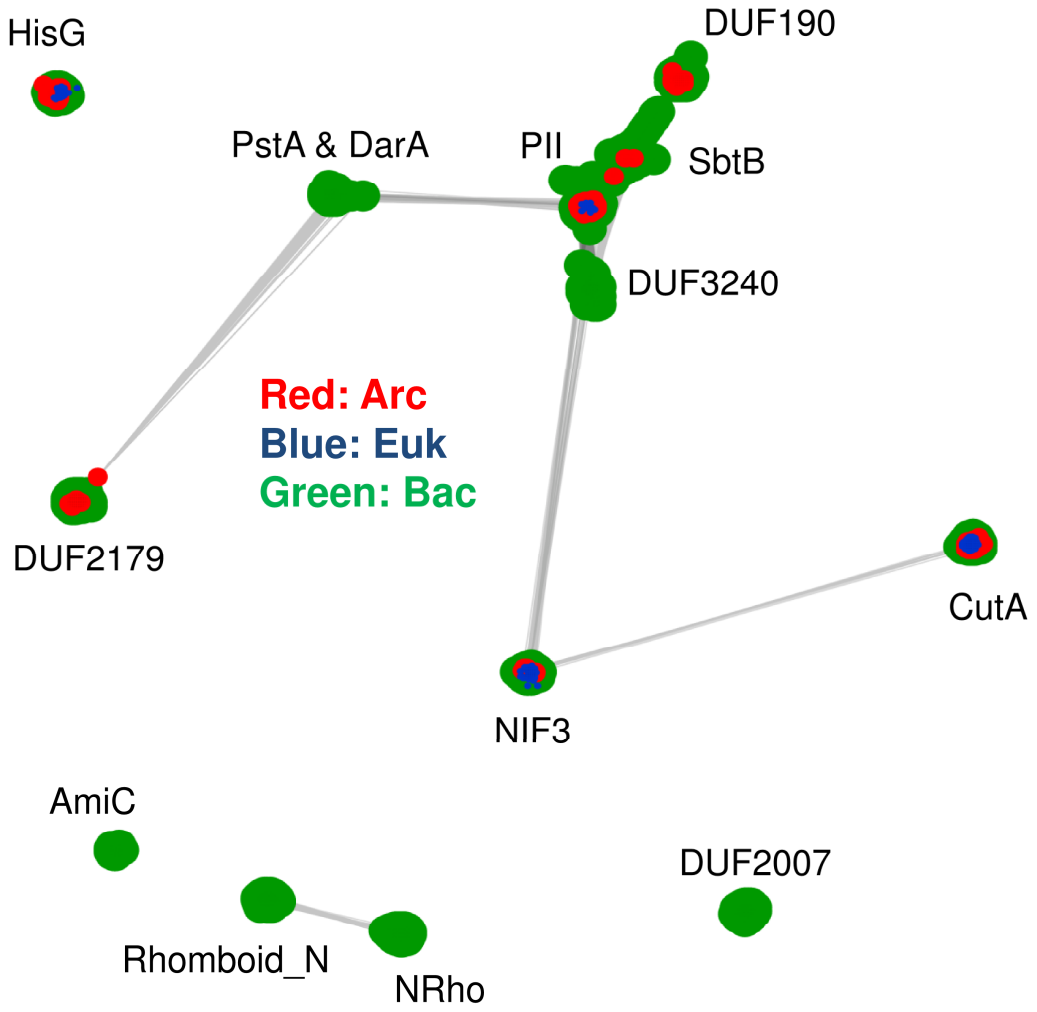
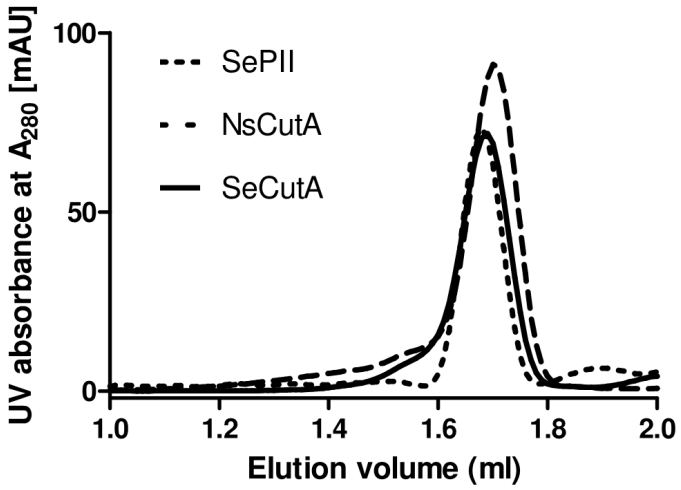


Fig. 6. Phenotypic characterization of knockdown and knockout mutants of CutA. (A & C) The AS_*Ns*CutA and ΔSeCutA strains were tested against elevated concentrations of Cu^{2+} and Pb^{2+} for 120 h, respectively. The survival of the strains in compare to the WT were evaluated by drop-assay on normal BG11. (B) Growth curve of AS_*Ns*CutA strain under no or 25 µM Cu^{2+}, reveals no difference to the WT. (D) Microscopic examination of AS_*Ns*CutA strain under no or 25 µM Cu^{2+}, reveals no difference to the WT. The scale bars represent 10 µm.

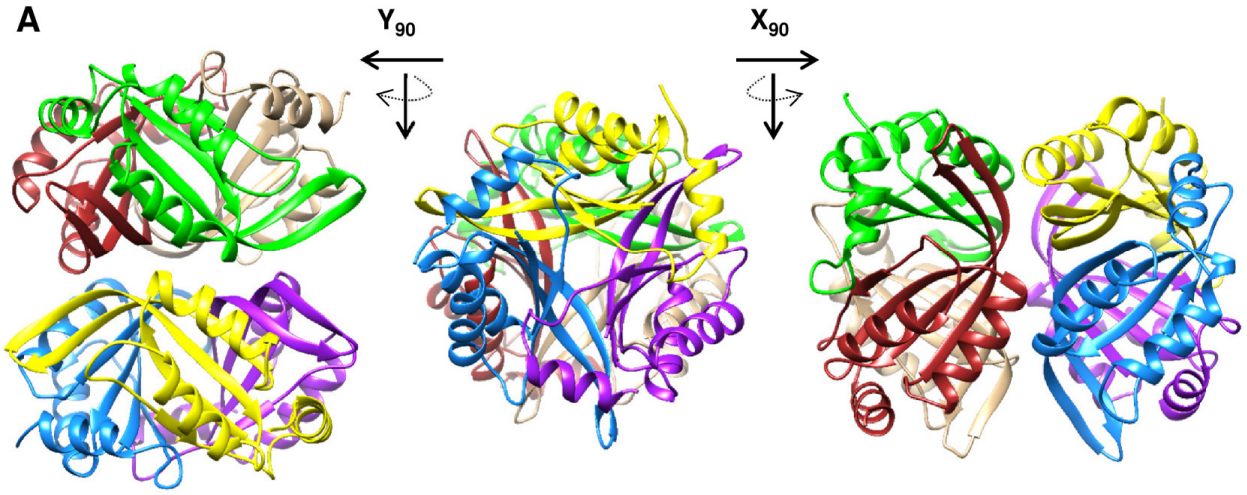
Supplementary Figures



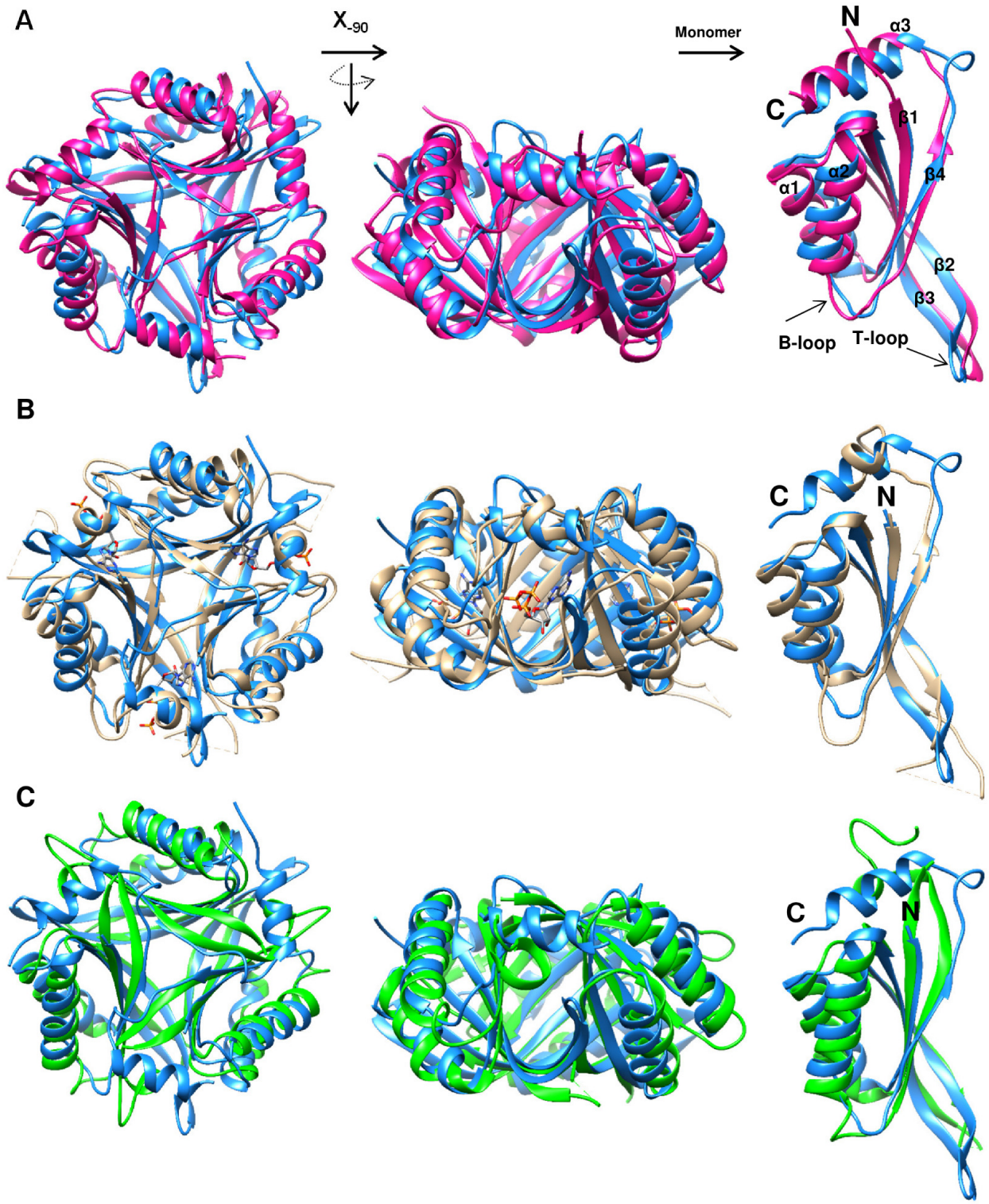
Supplementary Fig. S1. Cluster mapping showing the distribution of the members of PII superfamily across different domains of the life, as indicated.



Supplementary Fig. S2. Analytical gel filtration shows the elution profile of trimeric *Se*PII, *Ns*CutA, and *Se*CutA proteins.



Supplementary Fig. S3. Hexameric structures of *Ns*CutA as found in the crystal asymmetric unit represented as ribbons; each protomer is shown in a different color.



Supplementary Fig. S4. (A) Superposition of trimeric and monomeric subunit of *Ns*CutA (blue) over *Ec*CutA (pink). (B) Superposition of trimeric and monomeric subunit of *Ns*CutA (blue) over *Se*PII (brown). (C) Superposition of trimeric and monomeric subunit of *Ns*CutA (blue) over SbtB (green).

```
alr7093_Ns        ------------------------------------------------------------  0
Synpcc7942_Se     ------------------------------------------------------MTDLS  5
1VHF_Tm           ------------------------------------------------------------  0
1KR4_Tm           ---------------------------------------------MGSSHHHHHHSSGRE  15
4NYP_Ph           ------------------------------------------------------------  0
4NYO_Ph           ------------------------------------------------------------  0
1UMJ_Ph           ------------------------------------------------------------  0
3GSD_Yp           ------------------------------------------------SNAMSDSDAMTDP  13
3X3U_Ec           ---------------------------------------------------MLDE----K-  5
1NAQ_Ec           ---------------------------------------------------MLDE----K-  5
3OPK_Se_sub.      ------------------------------------------------SNAMLDV----KS  9
2NUH_Xf           ------------------------------------------------------------  0
3AHP_Ss           ------------------------------------------------------------  0
1P1L_Af           ------------------------------------------------------------  0
2ZFH_Hs           MSGGRAPAVLLGGVASLLLSFVWMPALLPVASRLLLLPRVLLTMASGSPPTQPSPASDSG  60
1XK8_Hs           --------------------------------MGGSHHHHHHGMASGSPPTQPSPASDSG  28
1OSC_Rat          ---------------------------------------------MASGSPPSQPSPA--SG  15
1V6H_Tt           ------------------------------------------------------------  0
1NZA_Tt           ------------------------------------------------------------  0
2ZOM_Os           -------------------------------------------------------MEST  4


                                                          38 40      47 49
alr7093_Ns        -------MKIALTNLP-PEHGERIARLLVEEHIVACVNLYP-VHSIYSWKGEVCSEAEVT  51
Synpcc7942_Se     SLIETADLRLLLTTVSTEVEAQQLAQAAVEAGLAACVSITP-IQSCYRWQGAIARETEQQ  64
1VHF_Tm           -----MSLILVYSTFPNEEKALEIGRKLLEKRLIACFNAFE-IRSGYWWKGEIVQDKEWA  54
1KR4_Tm           ALYFMGHMILVYSTFPNEEKALEIGRKLLEKRLIACFNAFE-IRSGYWWKGEIVQDKEWA  74
4NYP_Ph           -------MIIVYTTFPDWESAEKVVKTLLKERLIACANLRE-HRAFYWWEGKIEEDKEVG  52
4NYO_Ph           -------MIIVYTTFPDWESAEKVVKTLLKERLIACANLRE-HRAFYWWEGKIEEDKEVG  52
1UMJ_Ph           -------MIIVYTTFPDWESAEKVVKTLLKERLIACANLRE-HRAFYWWEGKIEEDKEVG  52
3GSD_Yp           NAVSYSNAIVVLCTAPDEASAQNLAAQVLGEKLAACVTLLPGATSLYYWEGKLEQEYEVQ  73
3X3U_Ec           --SSNTASVVVLCTAPDEATAQDLAAKVLAEKLAACATLIPGATSLYYWEGKLEQEYEVQ  63
1NAQ_Ec           --SSNTASVVVLCTAPDEATAQDLAAKVLAEKLAACATLIPGATSLYYWEGKLEQEYEVQ  63
3OPK_Se_sub.      QDISIPEAVVVLCTAPDEATAQDLAAKVLAEKLAACATLLPGATSLYYWEGKLEQEYEVQ  69
2NUH_Xf           ---MASDVYLIFSTCPDLPSAEIISRVLVQERLAACVTQLPGAVSTYRWQGKIETTQEIQ  57
3AHP_Ss           --MYKPEQLLIFTTCPDADIACRIATALVEAKLAACVQIGQAVESIYQWDNNICQSHEVP  58
1P1L_Af           ------MHNFIYITAPSLEEAERIAKRLLEKKLAACVNIFP-IKSFFWWEGKIEAATEFA  53
2ZFH_Hs           SGYVPGSVSAAFVTCPNEKVAKEIARAVVEKRLAACVNLIPQITSIYEWKGKIEEDSEVL  120
1XK8_Hs           SGYVPGSVSAAFVTCPNEKVAKEIARAVVEKRLAACVNLIPQITSIYEWKGKIEEDSEVL  88
1OSC_Rat          SGYVPGSVSAAFVTCPNEKVAKEIARAVVEKRLAACVNLIPQITSIYEWKGKIEEDSEVL  75
1V6H_Tt           ------MEEVVLITVPSEEVARTIAKALVEERLAACVNIVPGLTSIYRWQGEVVEDQELL  54
1NZA_Tt           ------MEEVVLITVPSEEVARTIAKALVEERLAACVNIVPGLTSIYRWQGEVVEDQELL  54
2ZOM_Os           STTV--PSIVVYVTVPNKEAGKRLAGSIISEKLAACVNIVPGIESVYWWEGKVQTDAEEL  62
                               .      .  :    :   : **        : : *.. :    *


                    55                   7475                 93 96
alr7093_Ns        LMMKVSTQGIERLKQRICELHPYELPEFVVIEVDNNASLREYIDFVKGETHLY--------  104
Synpcc7942_Se     MSFKTTVEQLDALQQWLQSQHPYALPECLVLTPIA--SSVAYRDWLRSSLS----------  113
1VHF_Tm           AIFKTTEEKEKELYEELRKLHPYETPAIFTLKVEN--VLTEYMNWLRESVLEGGS------  107
1KR4_Tm           AIFKTTEEKEKELYEELRKLHPYETPAIFTLKVEN--ILTEYMNWLRESVLGS--------  125
4NYP_Ph           AILKTREDLWEELKERIKELHPYDVPAIIRIDVDD--VNEDYLKWLIEETKK---------  102
4NYO_Ph           AILKTREDLWEELKERIKELHPYDVPAIIRIDVDD--VNEDYLKWLIEETKK---------  102
1UMJ_Ph           AILKTREDLWEELKERIKELHPYDVPAIIRIDVDD--VNEDYLKWLIEETKK---------  102
3GSD_Yp           LLFKSNTDHQQALLTYIKQHHPYQTPELLVLPVRD--GDKDYLSWLNASLL----------  122
3X3U_Ec           MILKTTVSHQQALLECLKSHHPYQTPELLVLPVTH--GDTDYLSWLNASLR----------  112
1NAQ_Ec           MILKTTVSHQQALLECLKSHHPYQTPELLVLPVTH--GDTDYLSWLNASLR----------  112
3OPK_Se_sub.      MILKTTVSHQQALIDCLKSHHPYQTPELLVLPVTH--GDTDYLSWLNASLR----------  118
2NUH_Xf           LLIKTNAVHVNAAITRLCALHPYRLPEAIAVQVSV--GLPEYLTWINTEIDEEYSLP----  112
3AHP_Ss           MQIKCMTTDYPAIEQLVITMHPYEVPEFIATPIIG--GFGPYLQWIKDNSPS---------  108
1P1L_Af           MIVKTRSEKFAEVRDEVKAMHSYTTPCICAIPIER--GLKEFLDWIDETVE----------  102
2ZFH_Hs           MMIKTQSSLVPALTDFVRSVHPYEVAEVIALPVEQ--GNFPYLQWVRQVTESVSDSITVLP  179
1XK8_Hs           MMIKTQSSLVPALTDFVRSVHPYEVAEVIALPVEQ--GNFPYLQWVRQVTESVSDSITVLP  147
1OSC_Rat          MMIKTQSSLVPALTEFVRSVHPYEVAEVIALPVEQ--GNPPYLHWVHQVTESV--------  126
1V6H_Tt           LLVKTTTHAFPKLKERVKALHPYTVPEIVALPIAE--GNREYLDWLRENTG----------  103
1NZA_Tt           LLVKTTTHAFPKLKERVKALHPYTVPEIVALPIAE--GNREYLDWLRENTG----------  103
2ZOM_Os           LIIKTRESLLDALTEHVKANHEYDVPEVIALPIKG--GNLKYLEWLKNSTRES--------  113
                    .*            :   * *                  :  ::
```

Supplementary Fig. S5. Multiple sequence alignment of different members of CutA.

Table (S1): Primers used in this study.

Primers/amplicon	**Name/Sequence 5'→3'**
Recombinant strep tagged CutA of *Nostoc* sp. PCC7120 (ORF *alr7093*)	1414: TGAATAGTTCGACAAAAATCTAGATAACGAGGGCAAAAAATGAAAATT GCATTAACAAATTTACCGCCAG
	1415: AAGCTTATTATTTTTCGAACTGCGGGTGGCTCCAAGCGCTATACAGAT GAGTTTCTCCTTTAACAAAGTC
Knockout construct of CutA in *Synechococcus elongatus* PCC 7942 (in ORF: *Synpcc7942_2261*)	
Kanamycin cassette	1238: GAACAGGCTTATGTCAATTCG
	1239: GAGTTTGTAGAAACGCAAAAAGGCCATCCGTC
Downstream of *Synpcc7942_2261*	1240: GTAAAACGACGGCCAGTGAATTCGAGCTCGGTACCCGGGGTCAGCGTG ATTGACCAAGCCAGCAGCACG
	1241: ATTACAGGGGTACCGAGCTCGAATTGACATAAGCCTGTTCAGCATCTG TTGCTTCTGATACATTG
Upstream of *Synpcc7942_2261*	1242: CCATCCTGACGGATGGCCTTTTTGCGTTTCTACAAACTCGGGCCCATT GAAGGCAGCGTCGGG
	1243: ATTACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCATGAC TGGACCGATCGCGACTG
Knockdown construct of CutA of ORF: *alr7093* in *Nostoc* sp. PCC7120	
Inducible promoter PetE	Fw: CTGCAGGTCGACTGCTAGAGGCATCAATTCGAGCTCGGTACCAGTACTCA GAATTTTTTGC
	1177: GGAGAAACTCATCTGTATTAAGGCGTTCTCCTAACCTGTAG
Reverse completed orientation of CutA (*alr7093*)	1178: CTACAGGTTAGGAGAACGCCTTAATACAGATGAGTTTCTCC
	1179: CTCATATGTATATCTCCTTCTTAAATCTAGAGGATCCCCGGATGAAAA TTGCATTAACAAA