The FEBS Journal

FEBS PRESS science publishing by scientists

# Functional and structural characterization of PII-like protein CutA does not support involvement in heavy metal tolerance and hints at a small-molecule carrying/signaling role

Khaled A. Selim[1,2], Lorena Tremiño[3], Clara Marco-Marín[3], Vikram Alva[2], Javier Espinosa[4], Asunción Contreras[4], Marcus D. Hartmann[2], Karl Forchhammer[1] and Vicente Rubio[3]

1 Interfaculty Institute for Microbiology and Infection Medicine, Organismic Interactions Department, Tübingen University, Germany
2 Department of Protein Evolution, Max Planck Institute for Developmental Biology, Tübingen, Germany
3 Instituto de Biomedicina de Valencia (IBV-CSIC), CIBER de Enfermedades Raras (CIBERER-ISCIII), Valencia, Spain
4 Departamento de Fisiología, Genética y Microbiología, Universidad de Alicante, Spain

**Keywords**
cyanobacteria; heavy metal tolerance; *Nostoc* sp. PCC 7120; PII superfamily; PII-like protein CutA; signal transduction; *Synechococcus elongatus* PCC 7942

**Correspondence**
K. A. Selim, Interfaculty Institute for Microbiology and Infection Medicine, Organismic Interactions Department, Tübingen University, Auf der Morgenstelle 28, 72076 Tübingen, Germany
Email: kKhaled.selim@uni-tuebingen.de
V. Rubio, Instituto de Biomedicina de Valencia (IBV-CSIC), and CIBER de Enfermedades Raras (CIBERER-ISCIII), Valencia, Spain
Email: rubio@ibv.csic.es

K. A. Selim and L. Tremiño are co-first authors

(Received 24 October 2019, revised 26 April 2020, accepted 1 June 2020)

doi:10.1111/febs.15464

The PII-like protein CutA is annotated as being involved in $Cu^{2+}$ tolerance, based on analysis of *Escherichia coli* mutants. However, the precise cellular function of CutA remains unclear. Our bioinformatic analysis reveals that CutA proteins are universally distributed across all domains of life. Based on sequence-based clustering, we chose representative cyanobacterial CutA proteins for physiological, biochemical, and structural characterization and examined their involvement in heavy metal tolerance, by generating CutA mutants in filamentous *Nostoc* sp. and in unicellular *Synechococcus elongatus*. However, we were unable to find any involvement of cyanobacterial CutA in metal tolerance under various conditions. This prompted us to re-examine experimentally the role of CutA in protecting *E. coli* from $Cu^{2+}$. Since we found no effect on copper tolerance, we conclude that CutA plays a different role that is not involved in metal protection. We resolved high-resolution CutA structures from *Nostoc* and *S. elongatus*. Similarly to their counterpart from *E. coli* and to canonical PII proteins, cyanobacterial CutA proteins are trimeric in solution and in crystal structure; however, no binding affinity for small signaling molecules or for $Cu^{2+}$ could be detected. The clefts between the CutA subunits, corresponding to the binding pockets of PII proteins, are formed by conserved aromatic and charged residues, suggesting a conserved binding/signaling function for CutA. In fact, we find binding of organic Bis-Tris/MES molecules in CutA crystal structures, revealing a strong tendency of these pockets to accommodate cargo. This highlights the need to search for the potential physiological ligands and for their signaling functions upon binding to CutA.

**Abbreviations**
*AS cutA*, *antisense cutA*-expressing strain of *Nostoc* sp. PCC7120; ASU and ASUs, singular and plural of asymmetric unit; *At*CutA, *Ec*CutA, *Ns*CutA, *Ph*CutA, *Rn*CutA, and *Se*CutA, CutA proteins from, respectively, *Arabidopsis thaliana*, *Escherichia coli*, *Nostoc* sp., *Pyrococcus horikoshi*, *Rattus norvegicus*, and *Synechococcus elongatus*; βA, β-amyloid; Bis-Tris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; B-loop and T-loop, characteristic loops found in the structure of canonical PII proteins; vestigial T-loop, small hairpin loop observed in CutA proteins instead of the T-loop of canonical PII; CPII, carboxysome-related PII protein; HEPES, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); IPTG, isopropyl β-D-1-thiogalactopiranoside; MES, 2-(N-morpholino)ethanesulfonic acid; *Nostoc*, *Nostoc* sp. PCC 7120; Ns1 and Ns2, two different crystals of the CutA protein from *Nostoc*; $OD^{xxx}$, absorbance at the wavelength indicated in the superindex, for 1-cm light path (unless indicated; PDB, Protein Data Bank; rmsd, root mean square deviation; *S. elongatus*, *Synechococcus elongatus* PCC 7942; SeL0, SeL1 SeL2 and SeL3 different crystals of the CutA protein from *S. elongatus* which host per trimer 0–3 molecules of bound Bis-Tris buffer; *Se*PII, canonical PII protein from *S.* elongatus; TLS, Translation/Libration/Screw.



The FEBS Journal **288** (2021) 1142–1162 © 2020 The Authors. *The FEBS Journal* published by John Wiley & Sons Ltd on behalf of Federation of European Biochemical Societies

**Databases**

Structural data are available in Protein Data Bank (PDB) under the accession numbers 6GDU, 6GDV, 6GDW, 6GDX, 6T76, and 6T7E.

## Introduction

PII proteins are widespread and ancient signaling molecules first reported in *Escherichia coli* as nitrogen metabolism controllers (reviewed in [1]). They are characteristic homotrimers presenting a $(\beta\alpha\beta)_2$ ferredoxin subunit fold (reviewed in [2,3]). Their targets are other protein molecules, either enzymes, transporters, or gene expression regulators (reviewed in [2,3]). Initially grouped into GlnK and GlnB classes depending on whether they were or were not involved in ammonia transport control [1,4], they initiated a superfamily that contains canonical PII proteins of known function as well as noncanonical members [5,6].

Despite the highly conserved structure, amino acid sequence conservation within the superfamily can be quite low [6,7]. The CutA protein family belongs to the families with the lowest sequence similarities to the canonical PII protein family. In fact, CutA was identified as a member of the PII superfamily because it was 'fished' in transitive BLAST searches using as baits the sequences of some Nif3 proteins (COG3323 entry in the EGGNOG database, http://eggnogdb.embl.de; they exhibit significant homology with PII [5]) and also because of the structural similarity of CutA with canonical PII [8]. However, the taxonomic distribution of CutA is even wider than that of canonical PII, which is not present in eukaryotes except Archaeplastida, suggesting an important role across phyla.

There is very little sound information on the functions of CutA proteins. Nevertheless, these proteins are widely annotated in databases as 'Divalent-cation tolerance protein CutA', and 'Involved in resistance toward heavy metals'. These annotations stem from the fact that the gene encoding CutA, called *cutA1*, was first identified in *Escherichia coli* in investigations of a locus, *cutA*, involved in divalent metal tolerance [9]. The *cutA* locus contains three genes: one encoding a small PII-like protein of ~ 13 kDa (CutA1) located in one operon, and two encoding inner membrane proteins (*cutA2-3*). Mutation in the *cutA* locus caused increased sensitivity of *E. coli* cells to divalent metal ions (copper, zinc, nickel, cobalt, and cadmium). Complementation of Δ*cutA* locus with the ORF encoding CutA1 conferred enhanced tolerance only toward elevated $Cd^{2+}$ levels, while it did not confer tolerance against $Cu^{2+}$, $Zn^{2+}$, or $Ni^{2+}$. Only complementation with both, *cutA1* and *cutA2* genes, restored the tolerance toward high $Cu^{2+}$ and $Ni^{2+}$ levels. From these data, the product of the *cutA1* gene was speculated to confer heavy metal tolerance or to be involved in divalent cation homeostasis [8,9], leading to the above-indicated annotation without further physiological examination in other bacterial species.

Because of the wide distribution of CutA proteins among species, the investigation of the functions of these proteins has considerable interest. For example, in humans, *cutA* mRNA (with several splicing forms [10]) and the protein itself are widely distributed through the body, with CutA protein levels being highest in the brain (https://bgee.org/?page = gene&gene_id = ENSG00000112514; https://www.proteomicsdb.org/proteomicsdb/#protein/proteinDetails/49653/expression), where CutA appeared to be involved in the processing, trafficking, and membrane-anchoring of acetylcholinesterase, although not via direct interactions with this key neurotransmitter-processing enzyme [10–13]. In addition, a membrane-anchored splicing isoform of CutA interacts with the membrane-bound β-secretase BACE1, which produces the β-amyloid (βA) neurotoxic peptide centrally linked to Alzheimer disease [14]. CutA knockdown or overexpression influenced in opposite directions βA secretion in an in vitro cellular system [14]. Indirect evidence has linked (reviewed in [15]) CutA to human reproduction and development, in particular to the growth and differentiation of embryonic stem cells, the survival and maturation of oocytes, the recognition of the gametes in the female genital apparatus, and the activation of secretion and development of the mammary gland. In fact, *cutA* has been identified [16] as one of the 35 key genes for milk secretion. The specific roles of CutA in all these processes are far from clear.

Although protein structures are believed to be crucial for deciphering the molecular bases of function [17], the present structural knowledge of CutA proteins does not parallel our (rather poor) understanding of the functions of these proteins. The first crystal structures of CutA proteins were determined for the bacterium *E. coli* (*Ec*CutA; PDB: 1NAQ) and the mammal *Rattus norvegicus* (*Rn*CutA; PDB: 1OSC) [8]. Both proteins formed trimers, indicating the evolutionary conservation of the trimeric architecture of PII-like proteins [8]. As already indicated, the subunits of the trimer exhibited a ferredoxin-like fold with a β1α1β2-β3α2β4 architecture. The fold included a small β-hairpin loop connecting strands β2 and β3, as well as an additional small C-terminal β5-α3 extension. In an *Ec*CutA crystal structure, a $Hg^{2+}$ metal ion was found bound within one of the intersubunit clefts [8], which was considered suggestive evidence for the involvement of CutA in heavy metal sensing.

To gain further insights into the distribution of CutA proteins among all domains of life, we performed a bioinformatic analysis to identify different subgroups of CutA proteins and their relation to the other members of the PII superfamily. In order to understand their function in cyanobacteria, we chose two cyanobacterial CutA proteins from *Synechococcus elongatus* PCC 7942 and *Nostoc* sp. PCC 7120 for physiological, biochemical, and structural characterization. In addition, we reexamined with up-to-date microbiological tools whether *cutA1* is or is not important for $Cu^{2+}$ resistance of *E. coli*.

# Results

## Cluster analysis of CutA proteins reveals widespread distribution among all domains of life

To ascertain the sequence conservation of CutA homologs across prokaryotes and eukaryotes, we carried out PSI-BLAST [18] searches seeded with representative CutA homologs. Our searches revealed that the CutA protein family is quite conserved, with most homologs sharing over 35% pairwise sequence identity; the human and *E. coli* CutA homologs, for instance, exhibit a pairwise sequence identity of ~ 43%. To detect other protein families homologous to CutA, we searched the PDB70 profile-HMM database using the remote homology detection method HHpred [19], with the *E. coli* CutA protein as query. 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 and ECOD protein classification databases [20].

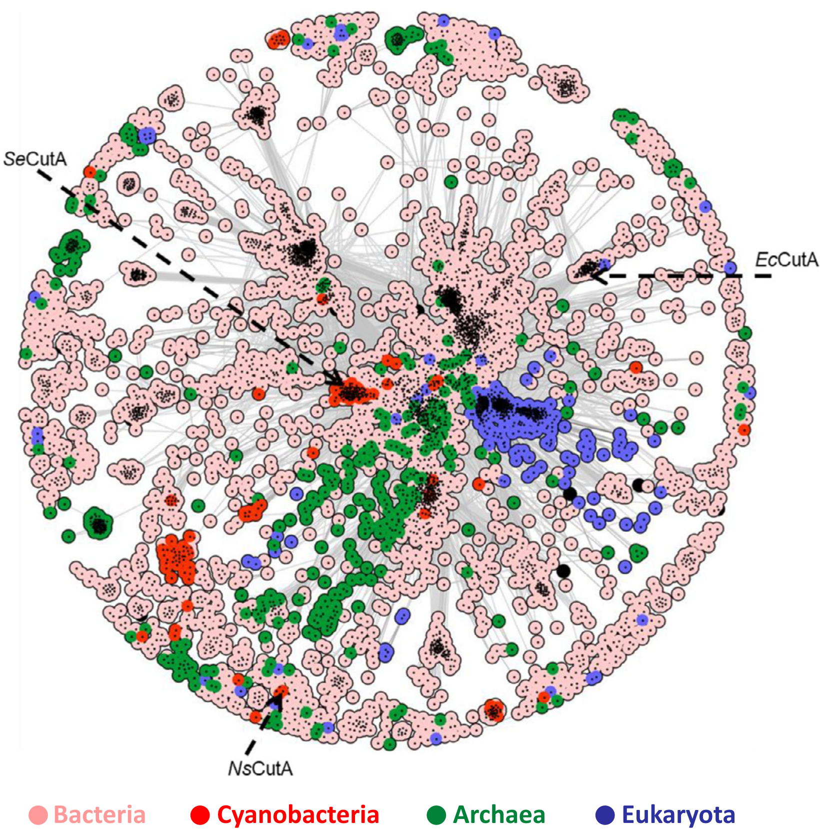



**Fig. 1.** Cluster map of CutA proteins. CutA proteins were gathered from the nr90 database using PSI-BLAST and were clustered in CLANS [21] based on the strength of their all-against-all pairwise sequence similarities. In the map, each dot represents one protein sequence and the sequences belonging to the same taxonomic group are shown in one color. BLAST connections are shown as gray lines to indicate the significance of sequence similarities; the darker a line, the higher the significance. Bacterial sequences are colored in light red, cyanobacterial sequences in red, archaeal sequences in green, and eukaryotic sequences in violet. The CutA proteins from *E. coli* (*Ec*CutA), *Nostoc* sp. PCC 7120 (*Ns*CutA), and *S. elongatus* PCC 7942 (*Se*CutA) are indicated by arrows.

To comprehensively explore the taxonomic distribution and evolutionary conservation of CutA proteins, we searched the nr90 database, a version of the nonredundant protein sequence database filtered to a maximum pairwise sequence identity of 90%, for CutA homologs using PSI-BLAST and investigated the obtained homologs using cluster analysis. The PSI-BLAST search yielded a total of 5736 homologs, which we subsequently clustered in CLANS [21] based on the statistical significance of their all-against-all pairwise sequence similarities, as measured by PSI-BLAST P-values. Since CutA proteins exhibit high pairwise sequence identities, we chose a stringent P-value cutoff to achieve separation between the archaeal, bacterial, and eukaryotic sequences (Fig. 1). At our chosen cutoff 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 eukaryotic CutA proteins are found primarily in alveolates, kinetoplastids, metazoans, and green plants, but are largely absent in fungi. By contrast, the prokaryotic CutA proteins 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 [22]. The cyanobacterial sequences (red) are mostly organized into two groups, but some are scattered across the map. While one of the two cyanobacterial groups is located near the eukaryotic cluster, the other is more distant.

To gain insights into the cellular function of CutA, in particular, the potential involvement of cyanobacterial CutA proteins in heavy metal tolerance, we chose two cyanobacterial homologs for biochemical and structural analysis. One of them, from the unicellular cyanobacterium *Synechococcus elongatus* PCC 7942 (*Se*CutA encoded by the ORF *Synpcc7942_2261*), is located close to the eukaryotic cluster (Fig. 1), whereas the other one, from the filamentous cyanobacterium *Nostoc* sp. PCC 7120 (*Ns*CutA encoded by the ORF *alr7093)*, is located at the periphery of the map (Fig. 1). We note that these proteins are the only CutA paralogs in these two organisms.

## CutA does not mediate heavy metal tolerance in cyanobacteria

The physiological importance and possible role of CutA in the sensitivity of cyanobacteria to divalent metal cations were investigated by gene inactivation in *S. elongatus* PCC 7942 and *Nostoc* sp. PCC 7120 (from now on *S. elongatus* and *Nostoc*, respectively). First, we inactivated *S. elongatus cutA* (*Synpcc7942_2261)* by allelic replacement with the kanamycin-resistant allele *cutA*::Tn5 from cosmid 8S34-E4 [23] (Fig. 2A). After transformation of *S. elongatus* with cosmid 8S34-E4, candidate clones ($Km^R$ $Cm^S$) were isolated and subjected to PCR analyses. Complete segregation of the inactive *cutA*::Tn5 alleles was easily achieved (Fig. 2B), confirming the prior finding [24] that *cutA* is not an essential gene in *S. elongatus*. An alternative approach also generating *cutA* null kanamycin-resistant mutants in *S. elongatus* based on a pUC19 plasmid derivative as previously described for *Synechocystis* sp. PCC 6803 [7] was also successful (Fig. 2C).

The sensitivity of *S. elongatus* wild-type and *cutA* mutant to transition metals was monitored using cells grown to mid-exponential phase, which were spotted on the surface of BG11 agar plates supplemented or not with the indicated metal cation (done with $Cu^{2+}$, $Zn^{2+}$ and $Co^{2+}$, Fig. 2D). The addition of these metals significantly impaired cell growth, judged from the cell biomass required in the inoculum to observe growth. No differences in the sensitivities to these metals were observed between the wild-type and the *cutA* mutant (Fig. 2D). The number of metal cations tested was extended to also include $Cr^{2+}$, $Cd^{2+}$, $Ni^{2+}$, $Mn^{2+}$, $Fe^{2+}$, and $Pb^{2+}$, using growth assays in 24-well plates in liquid BG11 medium containing variable concentrations (maximal range, 0–50 µM) of these cations. Almost under all tested conditions, the mutant behaved with no notable difference from the wild-type strain (not shown). The *cutA* mutant only showed some differential sensitivity against high concentration of $Pb^{2+}$ (Fig. 2D). In nature, $Pb^{2+}$ is normally one of the trace heavy metals and occurs at very low concentrations, suggesting that the observed phenotype of the *cutA* mutant is not relevant in a natural environment.

In the filamentous multicellular cyanobacteria *Nostoc* sp. PCC 7120, attempts to inactivate by allelic replacement of *alr7093* gene (encoding CutA) with a kanamycin resistance cassette were unsuccessful and the inactive alleles could not be completely segregated (Fig. 3A), suggesting that *cutA* might be essential for the filamentous lifestyle of *Nostoc*. Therefore, we resorted to a previously implemented antisense-RNA knockdown strategy [25,26] based on transformation of *Nostoc* sp. PCC 7120 with the promoter-less pAM1956 vector carrying in a reverse-complement orientation the gene to be knocked down (*alr7093*) preceded by a strong copper-inducible promoter ($P_{petE}$)

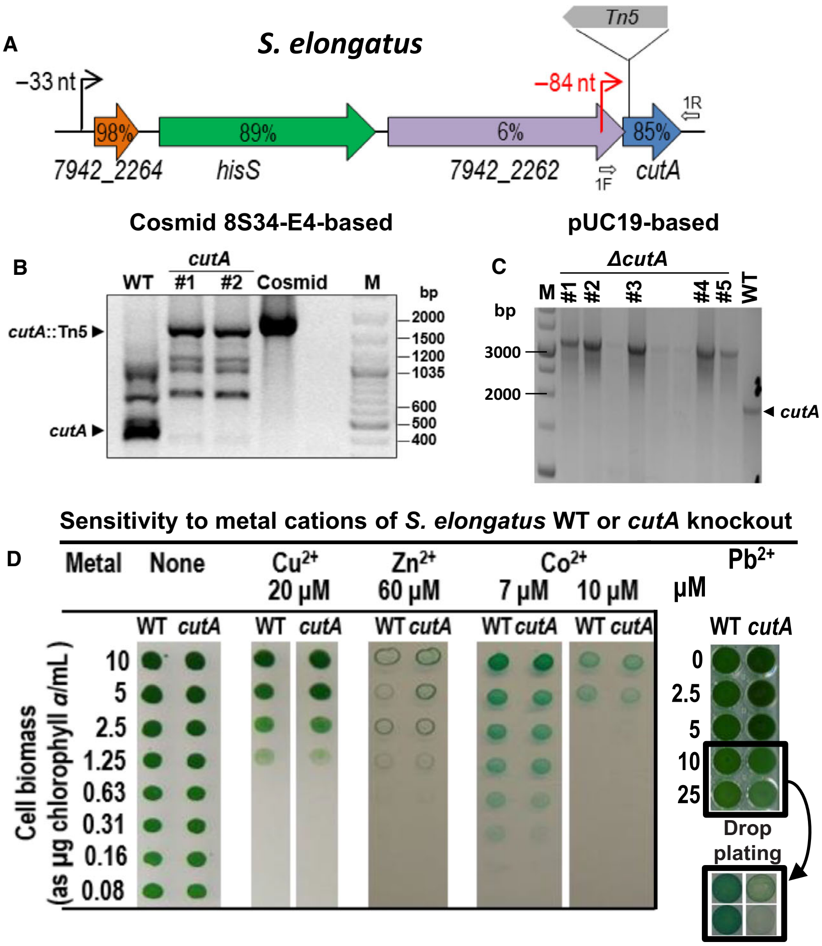



**Fig. 2.** Knockout *cutA* mutation does not alter metal sensitivity of unicellular cyanobacterium *S. elongatus* PCC 7942. (A) Schematic representation of the *Synpcc7942_2261* (designated *cutA*) genomic region in *S. elongatus* (shared by *Synechococcus elongatus* PCC 6301/UTEX2973, *Thermosynechococcus elongatus* BP-1, *Synechococcus* sp. PCC 7336, *Planktothrix agardhii* NIVA-CYA 126_8, *Synechococcus* sp. PCC 7003 and *Cyanothece* sp. PCC 7425). The conservation of each ORF (gene name or ORF ID) in cyanobacteria [73] is shown as percentage inside each ORF. Black and red arrows indicate reported [74,75] transcription start sites and distance to translation start sites for both 7942 and UTEX2973 (black) or just for UTEX2973 (red). The inactivation of the *cutA* gene was accomplished by Tn5 insertion of a kanamycin resistance cassette (the case shown, using cosmid-based inactivation) or by replacement of ORF *Synpcc7942_2261* by the cassette (pUC19-based inactivation). (B) Result of cosmid-based *cutA* inactivation. Two mutant *cutA* mutant clones (#1 and #2) alongside parental *S. elongatus* wild-type (WT) and the cosmid carrying *cutA*::Tn5 were PCR-analyzed with primers 1F and 1R. Detected alleles and reference size bands (bp) are indicated, respectively, on the left and the right of the gel shown at the bottom. (C) Electrophoretic analysis of PCR amplification of *cutA* reveals complete *cutA* inactivation, in the *S. elongatus* knockout mutant clones (#1 to #5) prepared by using a pUC19 plasmid derivative. (D) Wild-type and *cutA* mutant (cosmid-based) growth for 5 days on BG11 medium supplemented with the indicated divalent metal cations at the concentrations specified. For all metals except $Pb^{2+}$, serial dilutions of a cell culture (cell biomass assessed as chlorophyll α concentration in the culture) were spotted on the surface of agar plates of BG11 medium supplemented with the metal, which were photographed after 5 days of incubation at 28-30 °C under constant illumination conditions. In the case of $Pb^{2+}$, cells (wild-type and pUC19-based *cutA* mutant) were grown for 5 days in liquid BG11 medium supplemented with the indicated concentrations of $Pb^{2+}$, evaluating cell survival after 120 h for the highest $Pb^{2+}$ concentrations by drop-plating assay on normal BG11 medium.

and followed by a promoter-less reporter (*GFPmut2*) (Fig. 3B,C). The antisense *cutA*-expressing *Nostoc* strain (*AS cutA* strain) showed GFP fluorescence (Fig. 3D), reflecting expression of the downstream reporter, and thus of the *alr7093* antisense-RNA even at the low $Cu^{2+}$ concentrations (around 0.3 µM) present in the growth medium (BG11 medium). However, the *AS cutA* strain did not show increased metal sensitivity (tested for $Cu^{2+}$, $Fe^{2+}$, $Pb^{2+}$, $Co^{2+}$, $Zn^{2+}$, $Cr^{2+}$, $Cd^{2+}$, $Ni^{2+}$, and $Mn^{2+}$; illustrated only for $Cu^{2+}$ in

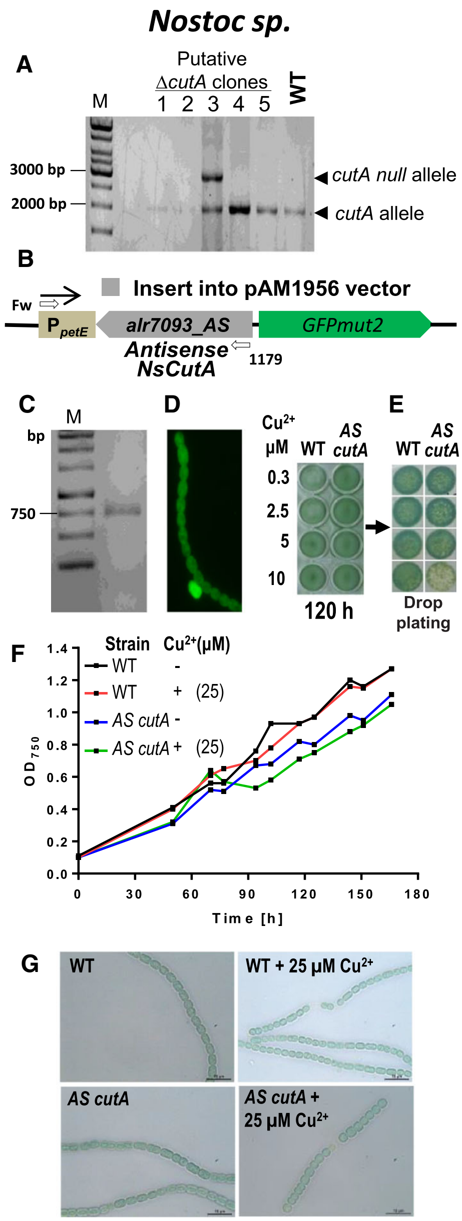



**Fig. 3.** *cutA* appears essential for *Nostoc* sp. PCC 7120, whereas the expression of antisense *cutA* does not substantially affect the metal sensitivity of *Nostoc*. (A) PCR analyses of 5 putative Δ*cutA* knockout clones obtained after transformation of wild-type *Nostoc* sp. (WT). Relevant size bands, from the marker (M lane), are indicated to the left and PCR-amplified alleles to the right. All tested clones retained wild-type *cutA* alleles. (B) Schematic representation of the *Antisense cutA* (*AS cutA*) into the pAM1956 vector inserted downstream of $Cu^{2+}$-inducible $P_{petE}$ promoter and followed by the GFP encoding gene (*GFPmut2*; in green). The arrow above indicates the direction of transcription from the $P_{petE}$ promoter. (C) PCR shows the presence of the *cutA-alr7093* gene in reverse-complement orientation under the control of $P_{petE}$ promoter within the *AS cutA* strain. The pair of primer Fw/1179 used for this PCR amplification is shown as empty arrows in (B). (D) Fluorescence microscopy of antisense knockdown strain reveals expression of GFP within the *AS cutA* strain, reporting the expression of *antisense cutA* in BG11 medium (containing 0.3 µM $Cu^{2+}$). (D) Growth for 5 days of wild-type or knocked down cells in liquid BG11 medium supplemented with the indicated concentrations of $Cu^{2+}$, evaluating cell survival after 120 h by drop assay on normal BG11 medium. (E) Growth curve of *Nostoc* WT and *AS cutA* strain in presence or absence of 25 µM $Cu^{2+}$ reveals no substantial differences for each strain between the $Cu^{2+}$-free and $Cu^{2+}$-containing cultures. (F) Illustrative images of the microscopic examination of *Nostoc* WT and *AS cutA* strains grown in the absence or in the presence of 25 µM $Cu^{2+}$. The scale bars represent 10 µm.

Fig. 3E), and the cells fully recovered as fast as the wild-type strain when returned to medium devoid of added metals (shown for $Cu^{2+}$ in Fig. 3E, *drop plating*). In addition, continuous growth monitoring in BG11 liquid medium (Fig. 3F) suggests that the *AS cutA* encoding plasmid might cause slight growth inhibition by itself, while no effect was observed by the addition of 25 µM $Cu^{2+}$ to the *AS cutA* strain, as it was also the case for the wild-type strain. Finally, microscopic examination (Fig. 3G) also did not reveal any morphological evidence for differential toxic effects between untreated and $Cu^{2+}$ (25 µM)-treated *AS cutA* cells, similarly to what was the case with wild-type cells. Overall, data on both *cutA* knockout in *S. elongatus* and *cutA* knockdown in *Nostoc* are consistent with CutA having little if any bearing on metal sensitivity of cyanobacteria under normal laboratory growth conditions.

## CutA does not mediate tolerance to $Cu^{2+}$ in *Escherichia coli*

The failure to substantially increase divalent metal sensitivity of two cyanobacteria by knockout or knockdown of *cutA* prompted us to re-examine the role of the corresponding *E. coli* gene, *cutA1*, since it was from data on this bacterium that the annotation of a role for this gene in divalent cation protection had stemmed [9]. We used strain JW4097 of the Keio collection [27], in which the *cutA1* gene of *E. coli* K-12 is replaced by a kanamycin resistance cassette (Fig. 4A). To enable the use of a pET15b expression plasmid carrying *cutA1* for complementation studies, we first lysogenized the JW4097 cells as well as the parental wild-type strain BW25113 [27] with phage λDE3, which carries the gene for T7 RNA polymerase. This lysogenization had no effect on the metal sensitivity of the wild-type or the JW4097 mutant strains (data not shown).

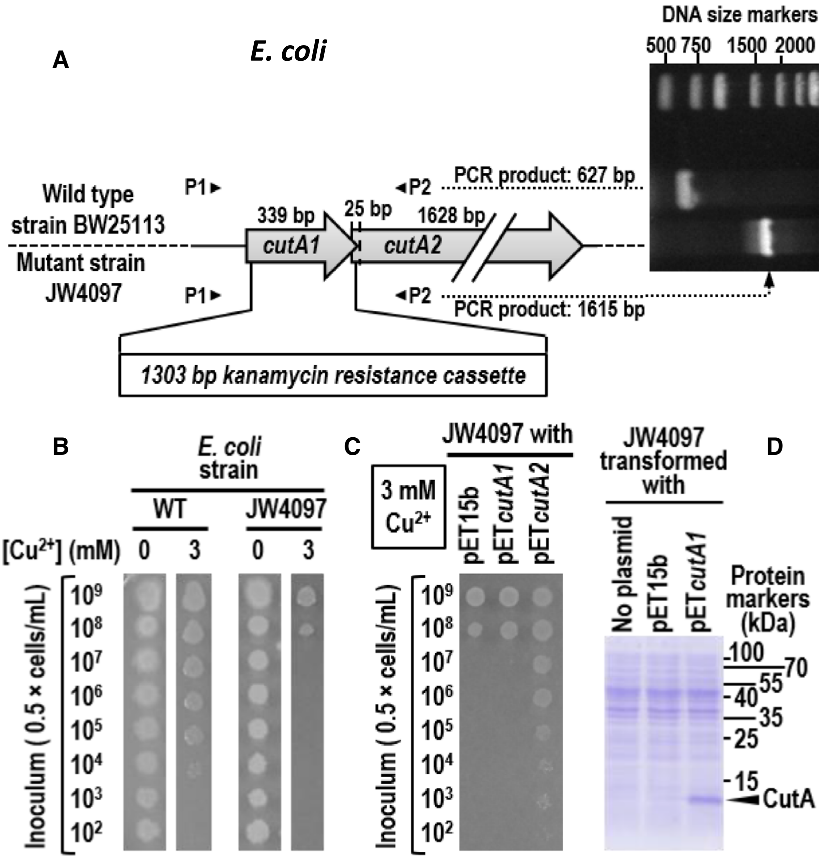
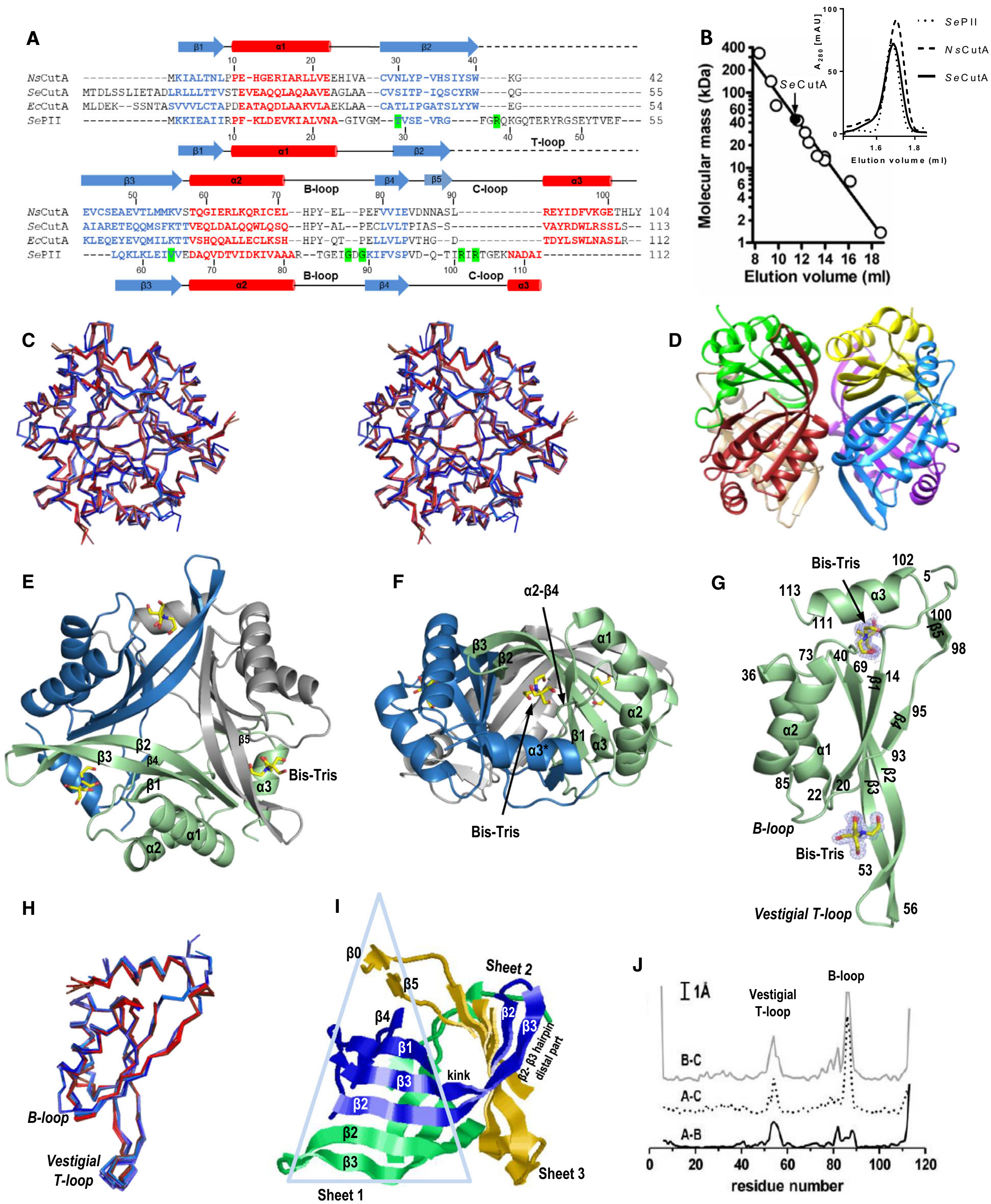



**Fig. 4.** Copper sensitivity of *E. coli cutA1* mutant and lack of complementation by transformation with *cutA1*-carrying plasmid. (A) *cutA1-cutA2* gene cluster of *E. coli K-12* in wild-type (WT) strain BW25113, and in mutant strain JW4097, in which *cutA1* is replaced by a kanamycin resistance cassette. To the right, electrophoretic analysis of the products obtained from the two strains after PCR amplification of the region encompassed between oligonucleotides P1 and P2. (B,C) Drop cultures of decreasing (from top to bottom) cell concentrations of both strains (lysogenized with λDE3 phage, see Materials and Methods) on agar-LB medium lacking or containing 3 mM copper sulfate as indicated. In (C), the cells were transformed with the indicated plasmid (see method section and text). (D) SDS/PAGE analysis (15% polyacrylamide gel) of $4 \times 10^8$ cells of *E. coli* JW4097 transformed as indicated. The gel was stained with Coomassie blue. The arrow marks the position of the CutA protein band.

The sensitivity of the JW4097 strain to $Cu^{2+}$ was considerably higher than that of the parental wild-type strain, as shown in drop cultures in which a $10^4$–$10^5$-fold higher cell inoculum was required with the JW4097 mutant than with the wild-type strain for growth in the presence of 3 mM $CuSO_4$ (Fig. 4B). However, this finding cannot be attributed conclusively to the lack of *cutA1*, since the initial 25 bases of the *cutA2* coding sequence (file P36655, UniProtKB database, https://www.uniprot.org/uniprot/P36655) overlap with the final bases of *cutA1* (Fig. 4A), and thus, a polar effect with lack of *cutA2* expression was anticipated in the mutant. Therefore, we examined separately the ability of *cutA1* and *cutA2* to complement this mutant strain in the copper sensitivity assays by transforming strain JW4097 with pET15 plasmids carrying either *cutA1* or *cutA2* (Fig. 4C). While resistance to $Cu^{2+}$ was restored to wild-type level by transformation with *cutA2* even in the absence of IPTG (the inducer of the T7 promoter; this finding indicates promoter leakiness), *cutA1* did not provide any growth advantage in the presence of 3 mM $Cu^{2+}$ over the same strain transformed with the empty parental plasmid (Fig. 4C). The lack of complementation by *cutA1* was not due to lack of *cutA1* expression, as proven by SDS/PAGE of cell extracts, which showed a band with the expected mass for CutA which was not present in the cells transformed with empty parental plasmid (Fig. 4D). Therefore, *E. coli* CutA does not provide protection against $Cu^{2+}$, indicating that CutA proteins from phylogenetically distant bacteria are not involved in resistance to heavy metal, the function previously attributed to CutA proteins.

## Structural characterization of *Se*CutA and *Ns*CutA

Although belonging to bacteria from the same taxonomic group (cyanobacteria), the CutA proteins from *S. elongatus*, *Se*CutA, and from *Nostoc*, *Ns*CutA, share only 26% sequence identity, and *Se*CutA is 12 residues longer at its N terminus than *Ns*CutA (Fig. 5A). In fact, the sequence of *Se*CutA resembles more in both length and identity (38%) the sequence of *E. coli* CutA (*Ec*CutA) than the one of *Ns*CutA (Fig. 5A). Thus, these two cyanobacterial CutA proteins could have important structural differences. This prompted us to explore these differences at the



structural level, while providing first examples of cyanobacterial CutA structures. With this goal, we produced *Se*CutA and *Ns*CutA in *E. coli* from their encoding genes carried in expression vectors. *Se*CutA was tagged either N-terminally with a poly-His tag (this was the form used here unless indicated), or with

**Fig. 5.** Cyanobacterial CutA architecture. Representations of protein structures were prepared using UCSF Chimera and PyMOL (see Materials and Methods). (A) Structure-based alignment of CutA proteins and of a canonical PII protein. Structures and sequences were retrieved from the indicated file identifiers (given between parentheses together with the abbreviation for the protein name) in the Protein Data Bank/UniProtKB databases (www.rcsb.org and www.uniprot.org/uniprot, respectively) for the CutA proteins from *Nostoc* sp. (*Ns*CutA; 6T76/**Q8YL42**), *S. elongatus* (*Se*CutA; 6GDX/**Q31KX8**) and *E. coli* (*Ec*CutA; 1NAQ/**P69488**; [8]), and for canonical PII from *S. elongatus* (*Se*PII; **2V5H/Q6V1L5**; [44]). The secondary structures of *Ns*CutA and *Se*PII proteins are shown above and below the alignment, respectively. The T-loops are indicated with dashed lines. The residues that mediate nucleotide binding in *Se*PII protein [28] are highlighted in green. Sequences were aligned with Clustal Omega, the alignment was manually refined taking in consideration the secondary structures, and it was represented using UCSF Chimera (see Materials and Methods). (B) Size-exclusion chromatography of *Se*CutA and *Ns*CutA. Main figure: semilogarithmic plot of molecular mass versus elution volume from a SuperdexTM 75 10/300 GL column of *Se*CutA and of protein standards. *Se*CutA is marked as a black circle, giving it the sequence-deduced mass for the homotrimer (~44 kDa). The standards used (open circles; their masses, in kDa, are given between parentheses) were as follows: bovine liver glutamate dehydrogenase (332.4), rabbit muscle lactate dehydrogenase (146), bovine serum albumin (66.4), maltose binding protein (42), bovine erythrocyte carbonic anhydrase (29), soybean trypsin inhibitor (21.5), cow's milk α-lactalbumin (14.2), bovine pancreas ribonuclease A (13.7), equine heart cytochrome C (12.4), bovine lung aprotinin (6.5), and vitamin B12 (1.35). Inset: elution from a Superdex$^{TM}$ 200 column PC 3.2/30 (GE Healthcare) of standard well-characterized Strep-tagged trimeric canonical *Se*PII [44], overlaid with the elution profiles for strep-tagged *Ns*CutA and *Se*CutA. (C) Stereo pairs of the superimposition of the main chain backbones of the structures of the four *Se*CutA trimers (different blue hues) and the eight *Ns*CutA trimers (different red hues) from the crystal structures reported here (Table 1). (D) Hexameric assembly in the asymmetric unit of crystal Ns1 (PDB: 6T76) in ribbon representation; each protomer is shown in a different color. (E, F) The CutA trimer found in crystal SeL3 (PDB: 6GDX), shown in cartoon representation with each subunit colored differently, in views (E) along the threefold axis or (F) with this axis vertical. Bound Bis-Tris molecules are shown in sticks representation, with C, O, and N atoms yellow, red, and blue, respectively. (G) The fold of one subunit of the trimer in the L3 crystal. The first and last residues of each secondary structure elements are given (by their numbering in the protein chain). The *2Fo–Fc* map for the two Bis-Tris molecules (labeled and represented in sticks representation) that interact with this subunit is contoured as a grid at σ = 1. (H) Superimpositions of the main chain of all the subunits in the ASUs of the different crystals (a different blue hue for each *Se*CutA crystal; a different red hue for crystals Ns1 and Ns2). (I) The β sheets that nucleate the trimer. One of the three sheets is enclosed in a blue triangle, with its β strands labeled. The kink in one of the β2β3 hairpins and the distal part of this hairpin are labeled. (J) Quantification of the distances between the Cα atoms at each position for the superimposition of the backbones of each possible pair of subunits in the SeL3 trimer.

a C-terminal Strep tag, while *Ns*CutA was Strep-tagged C-terminally (see Materials and Methods). These forms, purified with appropriate affinity columns, behaved as trimers in gel filtration, as shown by comparison with the elution of protein standards (Fig. 5B) or by co-elution with the canonical PII protein from *S. elongatus* PCC 7942 (*Se*PII, a very well-studied trimeric PII protein [2,3]) (Fig. 5B, inset).

We obtained crystals of *Se*CutA, called here SeL0, SeL1, SeL2, and SeL3, under four different conditions (Table 1). They diffracted X-rays at excellent resolutions (1.17–2.0 Å) and had similar monoclinic unit cells in space group $P2_1$. *Ns*CutA yielded an orthorhombic crystal form in $P2_12_12_1$ (Ns1, 1.9 Å resolution), and a trigonal one in P321 (Ns2, 2.45 Å). All structures could be solved by molecular replacement, and refinement yielded models with very good geometries (Table 1). The asymmetric units (ASUs) of the four *Se*CutA crystals contained identical single homotrimers (Fig. 5C; root mean square deviation (rmsd) values for superimposition of Cα atoms, 0.18–0.63 Å; mean, 0.48 Å) which hosted in each crystal a different number of bound molecules of Bis-Tris buffer (Table 1 and see below). Of note, the SeL2 crystal was grown in the presence of 0.22 mM $CuSO_4$ (Table 1) and yielded a model at excellent resolution (1.8 Å) that should have allowed the detection of bound $Cu^{2+}$, but none was observed, in agreement with the conclusion from our *in vivo* experiments that cyanobacterial CutA is not involved in protection against $Cu^{2+}$. Among the *Se*CutA crystals, SeL3 yielded the highest resolution (1.17 Å) structure ever reported for a CutA protein (Table S1). This crystal structure will be used with preference in further descriptions. Concerning the *Ns*CutA crystals, the ASU of Ns1 contained two identical homotrimers forming a bottom-to-bottom dimer of trimers (Fig. 5D), while the ASU of crystal Ns2 contained six CutA monomers. Each of these monomers in Ns2 belonged to a different trimer that was built by application of the crystal symmetry and that, by application of that symmetry, showed the same hexamerization behavior as in Ns1 (not shown). All trimers from Ns1 and those generated by application of crystal symmetry in Ns2 superimpose closely, with rmsd values ≤ 0.35 Å. In the case of the *Se*CutA crystals, no hexamers or higher oligomers than the trimer were observed or could be generated by crystal symmetry operations. The trimers found in the ASUs of the different *Se*CutA crystals superimposed quite well mutually and with those of *Ns*CutA (Fig. 5C, and

**Table 1.** Crystallization, X-ray data collection, and refinement statistics for the crystals of the two cyanobacterial CutA proteins studied here. Values in parenthesis are the data for the highest resolution shell

| Protein | *Se*CutA | | | | *Ns*CutA | |
|---|---|---|---|---|---|---|
| Crystal | SeL0[a] | SeL1 | SeL2 ($Cu^{2+}$) | SeL3 | Ns1 | Ns2 |
| PDB file | 6GDU | 6GDV | 6GDW | 6GDX | 6T76 | 6T7E |
| Reservoir solution | 0.1 M MIB[b] pH 4, 25% PEG1500 | 0.1 M Bis-Tris pH 5.5, 0.2 M $Li_2SO_4$, 25% PEG3350 | 0.1 M Bis-Tris pH 6.5, 29% PEG3350, 0.22 mM $CuSO_4$ | 0.1 M Bis-Tris pH 6.5, 23% PEG3350 | 2 M $(NH_4)_2SO_4$, 0.1 M Bis-Tris pH 6.5 | 0.1 M MES pH 6.5, 1.2 M K/Na tartrate |
| Data collection | | | | | | |
| Synchrotron/Beamline/ Detector | ALBA/BL13/ Dectris Pilatus 6M | | Diamond/I03/ PSI Pilatus 6M | ALBA/BL13/ Dectris Pilatus 6M | Swiss Light Source/X10SA/ Pilatus 6M-F | |
| Wavelength (Å) | 0.9794 | 0.9794 | 0.9795 | 0.9794 | 1.0 | 1.0 |
| Space group | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_12_12_1$ | P321 |
| Unit cell parameters | | | | | | |
| a, b, c (Å) | 35.2, 96.4, 49.0 | 34.6, 96.6, 48.9 | 34.7, 95.8, 48.8 | 35.6, 96.3, 48.9 | 62.1,90.7,93.7 | 87.1, 87.1, 153.0 |
| α, β, γ (°) | 90, 94.2, 90 | 90, 97.4, 90 | 90, 97.1, 90 | 90, 93.2, 90 | 90, 90, 90 | 90, 90, 120 |
| Solvent (%) | 42 | 45 | 42 | 44 | 35 | 49 |
| Resolution range[c] (Å) | 48.19–1.75 (1.78–1.75) | 48.32–2.00 (2.05–2.00) | 47.91–1.80 (1.85–1.80) | 48.84 –1.17 (1.19–1.17) | 46.85–1.90 (2.01–1.90) | 38.25–2.45 (2.60–2.45) |
| Unique reflections[c] | 32,752 (1,772) | 20,245 (1,498) | 28,906 (2,129) | 110,414 (5,494) | 42,415(6,637) | 25,325 (3,985) |
| Completeness[c] (%) | 99.9 (100.0) | 94.2 (93.2) | 98.7 (98.8) | 99.9 (99.6) | 99.7 (97.9) | 99.8 (99.4) |
| Redundancy[c] | 7.2 (7.5) | 2.7 (2.8) | 3.2 (3.2) | 4.6 (3.9) | 6.6 (6.3) | 8.7 (9.0) |
| $I/\sigma$[c,d] | 11.2 (3.2) | 5.6 (1.6) | 10.7 (1.8) | 21.6 (5.3) | 13.4 (1.4) | 17.0 (2.6) |
| *R* values[c,d] (%) | $R_{pim}$ 4.1 (21.8) | $R_{pim}$ 7.6 (23.5) | $R_{meas}$ 10.5 (96.1) | $R_{pim}$ 1.7 (13.0) | $R_{merge}$ 9.9 (130.1) | $R_{merge}$ 9.2 (79.7) |
| CC(1/2)[c,d] | – | – | – | – | 99.9 | 99.9 (84.2) |
| Refinement | | | | | | |
| Resolution[c] (Å) | 48.19–1.75 (1.87–1.75) | 48.32–2.00 (2.05–2.00) | 47.91–1.80 (1.85 –1.80) | 48.84–1.17 (1.19–1.17) | 46.85–1.90 (1.95–1.90) | 38.25–2.45 (2.51–2.45) |
| Reflections, work/test | 31,114/1368 | 19,233/1012 | 27,461/1445 | 104,893/5521 | 40,283/2121 | 24,058/1267 |
| $R_{factor}/R_{free}$[e] (%) | 19.9/23.3 | 19.2/25.5 | 20.3/24.4 | 14.0/16.3 | 19.85/22.58 | 22.15/27.13 |
| rmsd from ideal | | | | | | |
| Bond length (Å) | 0.011 | 0.013 | 0.014 | 0.014 | 0.017 | 0.007 |
| Bond angle (°) | 1.457 | 1.570 | 1.646 | 1.646 | 1.801 | 1.243 |
| Number of | | | | | | |
| Polypeptide chains | 3 | 3 | 3 | 3 | 6 | 6 |
| Protein atoms | 2462 | 2521 | 2447 | 2554 | 4901 | 4874 |
| Water molecules | 122 | 135 | 131 | 366 | 102 | 16 |
| Ligand molecules | 0 | 1 Bis-Tris[f] | 2 Bis-Tris[f] | 3 Bis-Tris[f] | 7 $SO_4^{2-}$ ions | 4 MES[g] |
| Average B-factor (Å$^2$) | | | | | | |
| Protein | 15.5 | 25.2 | 31.6 | 14.9 | 38.5 | 61.7 |

**Table 1.** (Continued).

| Protein | *Se*CutA | | | | *Ns*CutA | |
|---|---|---|---|---|---|---|
| Crystal | SeL0[a] | SeL1 | SeL2 ($Cu^{2+}$) | SeL3 | Ns1 | Ns2 |
| Water molecules | 33.4 | 33.3 | 42.5 | 30.8 | 40.8 | 43.0 |
| Ligand molecules | - | 27.4 | 48.0 | 17.9 | 60.8 | 86.5 |
| Ramachandran[h] (%) | | | | | | |
| Favored | 100.0 | 100.0 | 100.0 | 100.0 | 94.9 | 95.3 |
| Allowed | 0.0 | 0.0 | 0.0 | 0.0 | 5.1 | 4.7 |
| Outliers | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |

[a]The protein solution was supplemented with 10 mm ATP and 10 mm $MgCl_2$.; [b]MIB is a commercial proprietary buffer from Molecular Dimensions, which contains disodium malonate, imidazole, and boric acid; PEG, polyethylene glycol.; [c]Values between parentheses correspond to the highest resolution shell.; [d]http://shelx.uni-ac.gwdg.de/~athorn/pdf/thorn_cshl2014_quality_indicators.pdf.; [e]https://strucbio.biologie.uni-konstanz.de/ccp4wiki/index.php/R-factors.; [f]Bis-Tris: Bis-(2-hydroxy-ethyl)-amino-tris(hydroxymethyl)-methane.; [g]MES: 2-(N-morpholino) ethanesulfonic acid.; [h]Calculated for *Se*CutA models with RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) and for *Ns*CutA models with PROCHECK (https://www.ebi.ac.uk/thornton-srv/software/PROCHECK).

rmsd values for trimer–trimer Cα-superimpositions in the range 1.20–1.37 Å; mean for 32 superimpositions, 1.27 Å), also exhibiting virtually perfect 3-fold molecular symmetry. Given the elution of *Ns*CutA as a trimer in size-exclusion chromatography (Fig. 5B) and the low energy (7 kcal·$mol^{-1}$) predicted to be needed for hexamer dissociation to trimers (PISA server prediction, https://www.ebi.ac.uk/pdbe/prot_int/pistart.html), a physiological occurrence of the hexamer appears unlikely. Thus, the hexameric assembly will not be regarded further.

The *Se*CutA and *Ns*CutA homotrimers (Fig. 5C, E and F) conform with the architecture of PII superfamily proteins (Fig. S1), which are homotrimers of subunits folded as a ferredoxin-like core followed by a C-terminal extension (Fig. 5G,H). In the core, the two interlocking βαβ motifs form a four-strand antiparallel β-sheet, whereas the two antiparallel helices sit on the sheet's outer surface (Fig. 5E,F). The folding of the C-terminal extension, somewhat variable among different members of the PII superfamily (Fig. S1), includes in CutA a beta strand (β5) followed by an alpha helix (α3) (Fig. 5G,H). A particular structural feature of *Se*CutA that has not been reported in other CutA proteins is the folding of its 12-residue N-terminal extension (the part missing in *Ns*CutA (Fig. 5A), but present although not visible in the structure of *Ec*CutA [8]) as an extra beta strand, called here β0, which interacts in an antiparallel manner with β5 from the same subunit, (Fig. 5I) perhaps helping stabilize the trimeric fold of *Se*CutA (see below).

In canonical PII proteins, the two βαβ motifs of the ferredoxin fold are connected by the long, surface-exposed, T-loop (Fig. S1B). Instead, in CutA proteins including *Se*CutA and *Ns*CutA, strands β2 and β3 are very long and are connected by a very short tip loop called here *vestigial T-loop* (Fig. 5G,I). These strands form a β hairpin that is kinked at the middle, changing its direction by nearly 90°. The distal part of this hairpin forms one edge of a central hybrid molecular antiparallel β sheet (Fig. 5I) composed of eight β strands with the topology $\beta3^{D}\beta2^{D}\mathit{\beta2\beta3\beta1\beta4}\underline{\beta5(\beta0)}$, where the four strands of the ferredoxin fold of one subunit (in italic lettering) form the center of the sheet, the distal part of the β2β3 hairpin (marked with the superscript D) of the preceding subunit (in the view of Fig. 5E) forms one edge of the sheet, and β5β0 (underlined; β0 is in parentheses because it exists only in *Se*CutA) of the third subunit form the other edge of the sheet (Fig. 5I). Because of the participation of the β2β3 hairpin of each subunit in two adjacent sheets, the three sheets that nucleate the trimer become continuous (Fig. 5I). Of note, the β4 strand and the preceding B-loop that connects it to helix α2 are shorter in CutA than in canonical PII (Fig. 5A and Fig. S1B), and helix α3, belonging to the C-terminal part of the polypeptide, is an important element in defining the intersubunit crevice (see below) (Fig. 5F).

Although in the trimers observed in the different crystals the three subunits were practically identical, some conformational differences were identified among subunits (Fig. 5H). These changes, best represented in the high-resolution structure of *Se*CutA (crystal SeL3), occur in two parts of the protein chain, in particular residues (*Se*CutA numbering) 52–57, in the vestigial *T*-loop, and at residues 81–89, in the B-loop (Fig. 5H,J), particularly at Pro86 and Tyr87 (respective displacements of their Cα atoms of up to 6.1Å and 4.8 Å, Fig. 5J). When residues 52–57 and 81–89 were excluded from the superimpositions, rmsd values for the three subunits of the trimer found in SeL3 were ≤ 0.54Å (93 Cα atoms superimposed), clearly illustrating the high structural stability of each subunit.

### Intersubunit cargo-carrying pockets

CutA proteins exhibit three pockets at intersubunit junctions, where the canonical PII proteins have their nucleotide-binding pockets [3,6,8,17]. Each pocket, reflecting the crevice between adjacent molecular β sheets, is delimited (Fig. 5E,F and Fig. 6A) by the arched β2β3 hairpin of a subunit and the C-terminal helix (helix α3) of the next subunit. In addition, the vestigial *T*-loop (β2-β3 connector), B-loop (α2-β4 connection), and the β1-α1 connector of the first of these subunits also contribute to delineate the pocket. The *Ns*CutA and *Se*CutA pockets are negatively charged (Fig. 6B,C) and highly polar, including highly conserved residues Tyr51, Glu60, Glu62, Tyr87, Glu91, and Tyr104 (*Se*CutA numbering used from here on; see Fig. 5A for correspondences with *Ns*CutA residues) (Fig. 6D,E). These residues largely define an opening (diameter of about 6 Å) deep in the pocket that connects it with the water-filled internal cavity of the trimer (not shown). Cys41 and Lys68 (the only positive residue in the pocket) are also constant and sit near this narrow opening (Fig. 6D). The conservation of these and other residues that form the pocket (Fig. 6E, Fig. S2) suggests important roles for these residues and for the pockets. The negative potential, the pocket size, and the lack of conservation of most ATP/ADP binding residues of canonical PII proteins (Fig. 5A) [28] seem to exclude the binding of these PII effectors to the CutA pockets. Indeed, isothermal titration calorimetry (ITC) failed to reveal any binding of ATP and ADP to *Ns*CutA (data not shown), which are effectors of canonical PII proteins [2,28], and of cAMP, which is the effector of the noncanonical PII-like protein SbtB [7]. Interestingly, in line with the results of our functional *in vivo* studies, we could not detect also the binding of $Cu^{2+}$ to recombinant CutA protein using ITC, among other tested metal cations. In addition, we could not detect any electron density for ATP or for $Mg^{2+}$ within the SeL0 crystal (1.75Å resolution), which grew in presence of ATP and $Mg^{2+}$ (10 mM each; Table 1).

The CutA pockets have been proposed to be functional elements that could bind some unrecognized ligand, playing crucial but as yet non-demonstrated CutA function(s) [17]. In agreement with this proposal, the SeL1, SeL2, and SeL3 crystals were found to host in these pockets, respectively, one, two, and three molecules of the Bis-Tris buffer present in the crystallization solutions (shown for SeL3 in Fig. 5E,F and Fig. 6C). Similarly, four of the six subunits in the ASU of the Ns2 crystal hosted a molecule of the MES buffer used in the Ns2 crystallization solution (Fig. 6A,B). The high-resolution structures of Bis-Tris bound to the *Se*CutA pockets showed (Fig. 6F) the buffer molecule lying on Trp107*, Tyr51, and Ala 40* (the asterisk marks elements from a second subunit) and hydrogen bonding, via its OH groups, to Leu 89, the side-chain carboxylate of Glu62, and, indirectly via a water molecule, to Cys41. Bis-Tris was positioned identically in two pockets of SeL3, while in the third pocket it was in an inverted orientation so that its ternary N and quaternary C atoms revert their respective positions relative to the other orientation. This allows preservation of most contacts, although being mediated by different OH groups of Bis-Tris, and adapts the ligand binding to movement of the B-loop (not shown). MES buffer binds to the pockets in Ns2 so that its morpholino part grossly occupies the position of Bis-Tris in *Se*CutA, while the negatively charged sulfate stems out from the site (Fig. 6A, B and G), as previously observed in the structure of CutA from *Pyrococcus horikoshi* (PDB: 4NYO) [29]. The same can be said for HEPES buffer, observed binding in the deposited structure (PDB: 3GSD; no publication associated) of CutA from *Yersinia pestis*, where the sulfonic acid also stems out and the remainder of the molecule replaces the Bis-Tris of *Se*CutA.

In the cyanobacterial structures, Bis-Tris or MES binding did not trigger any chain movements, as shown by superimposition of a chain with Bis-Tris of the SeL3 crystal and another one without Bis-Tris from SeL0. In the latter, structural differences occurred in the same regions and with similar

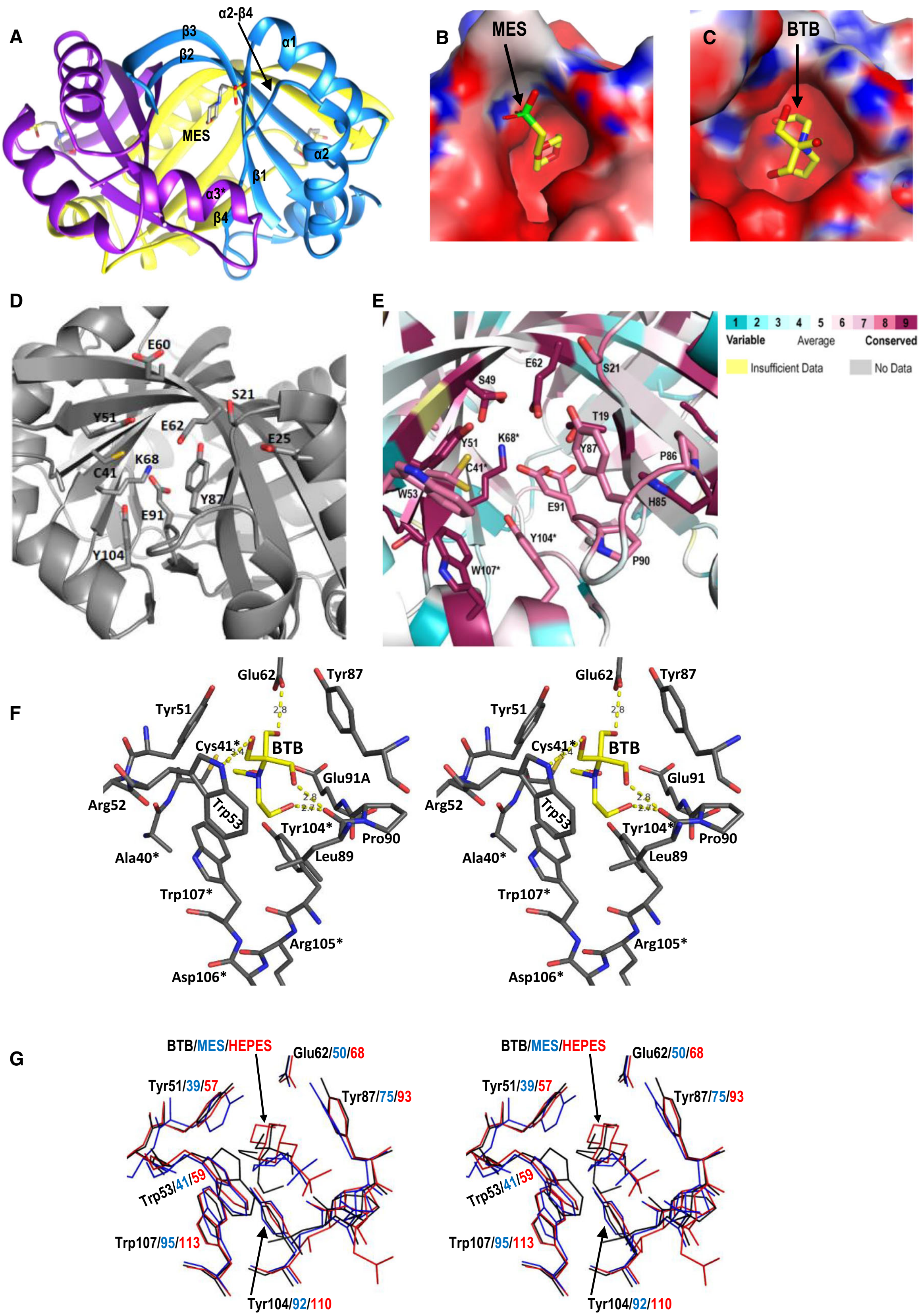

A
B
C
D
E
F
G
MES
BTB
α2-β4
β3
β2
α1
α2
β1
α3*
β4
E60
S21
Y51
E62
E25
K68
C41
Y87
E91
Y104
Variable
Average
Conserved
Insufficient Data
No Data
Glu62
Tyr87
Tyr51
Cys41*
Glu91A
Glu91
Arg52
Trp53
Tyr104*
Pro90
Ala40*
Leu89
Trp107*
Arg105*
Asp106*
BTB/MES/HEPES
Glu62/50/68
Tyr51/39/57
Tyr87/75/93
Trp53/41/59
Trp107/95/113
Tyr104/92/110

**Fig. 6.** Cargo-carrying pockets in cyanobacterial CutA proteins. Structural representations were prepared using UCSF Chimera and PyMOL (see Materials and Methods). (A) Ribbon representation of a trimer in the Ns2 crystal, to show the pocket and the bound MES molecule. Each subunit is represented in a different color. Some structural elements that conform the pocket are indicated, marking with an asterisk to differentiate the element from one subunit from those from the other subunit. (B) One pocket from the Ns2 crystal in surface view, with bound MES in sticks representation. The protein surface is colored according to its electrostatic surface potential, calculated with the programs PyMOL and APBS [76]: red–white–blue color gradient for surface potential values ranging from −5 to + 5 $k_BT/e$. (C) Same as (B) for *Se*CutA (SeL3 crystal) with bound Bis-Tris. (D,E) Cartoon representations of one pocket of the trimer in the SeL3 crystal showing (D) the richness of polar and negatively charged residues and illustrating the invariant Cys41 and the only positively charged residue (Lys68) of the pocket, while in (H) residue conservation (estimated by automatic alignment with the ConSurf Server; https://consurf.tau.ac.il [77]) is graded as indicated by color code. Residues from the subunit to the left are marked with an asterisk. (F) Respective stereo view of the sites formed by subunits C + A and A + B (SeL3 crystal trimer). Broken yellow lines, hydrogen bonds (distances in Å). Asterisks mark subunit C residues. O, N, and S atoms are colored red, blue, and gold, respectively, and C atoms are gray for the protein and yellow for Bis-Tris (abbreviated BTB). (G) Stereo view of the superimposition of the site shown in (F) with those for MES in a trimer of the Ns2 crystal structure and for HEPES in the deposited structure of *Yersinia pestis* (PDB: 3GSD).

magnitude (not shown) as among the subunits of the SeL3 trimer. Thus, mobility of the main chain in these regions appears not to be influenced notably by the bound Bis-Tris. A similar inference was made for the Ns2 crystal, since the MES-bound trimers superimpose with Cα rmsd values below 0.35 Å on the ligand-free trimers of the Ns1 crystal. Therefore, if Bis-Tris and MES mimic a physiological CutA ligand, their binding does not appear to induce conformational changes.

## Discussion

PII-like proteins are structurally similar and clearly related to canonical PII proteins, but they lack the PII PROSITE signature sequences, and their functions, when known, differ from those of classical PII proteins [6,7,30,31]. Bioinformatics and structural genomics approaches showed that PII-like proteins represent an even more widespread family of trimeric regulators than canonical PII proteins, distributed in almost all living organisms [6,30]. Among noncanonical members of the superfamily, a unique class of PII-like proteins is involved in controlling carbon metabolism, as exemplified by the carboxysome-related PII protein (CPII), which binds ADP/AMP and bicarbonate and was proposed to sense bicarbonate availability [31]. Some of us also characterized recently in cyanobacteria a new PII-like protein named SbtB [7], which senses the second messenger cAMP (in addition to ATP, ADP, and AMP) and which, being involved in the control of the bicarbonate transporter SbtA, links cAMP sensing to $CO_2$ metabolism [7,32]. Furthermore, in Firmicutes such as *Staphylococcus aureus* or *Bacillus subtilis*, a PII-like protein (termed PstA or DarA) was identified to sense the second messenger c-di-AMP, although its function is still unknown [33–35]. In all these cases, the PII/PII-like proteins exert their regulatory function through binding small effector molecules, mainly adenylyl nucleotides, which induces conformational changes to transduce the signal to the PII/PII-like targets [32]. We now deal with another noncanonical PII protein of unclear function, CutA, a protein that is distant in sequence from canonical PII proteins [5] but whose core structure is almost identical to the other PII superfamily members. However, it differs in the effector-binding site and the output domain, the T-loop, a minimal rudiment of the T-loop (termed vestigal T-loop) in CutA (Fig. S1B). These features set CutA apart from the other so-far characterized members of the PII superfamily. However, we could identify the binding pockets of *Se*CutA and *Ns*CutA, which hosted, respectively, Bis-Tris and MES molecules from the crystallization buffer, at the intersubunit clefts of the trimer. The fact that these pockets are lined by highly conserved residues among CutA proteins that differ from those residues used for the binding of nucleotides in characterized members of the PII protein superfamily signals the importance of these pockets. The respective binding of Bis-Tris and MES molecules to *Se*CutA and *Ns*CutA implies a tendency of CutA proteins to harbor small organic molecules with polar groups in these crevices, also observed in previous CutA structures from *Pyrococcus horikoshii* (PDB: 4NYO) and *Yersinia pestis* (PDB: 3GSD) bound to MES and HEPES, respectively. However, the natural CutA ligands of these sites as well as the conformational changes potentially triggered remain to be identified.

To our surprise, we were neither able to reveal binding of $Cu^{2+}$ to cyanobacterial CutA proteins *in vitro* by X-ray crystallography nor by ITC and could not show involvement of CutA in heavy metal tolerance *in vivo* by physiological means in three phylogenetically distant bacteria. Actually, $Cu^{2+}$ was previously observed [36] only in the structure of one CutA protein (PDB file 1UKU; for *P. horikoshi* CutA, *Ph*CutA) among the 18 crystal structures of CutA proteins reported or deposited in the Protein Data Bank (PDB)

(Table S1) belonging to 14 different organisms (two archaea, eight bacteria and four eukaryota, two of them mammals). Even for *Ph*CutA, later crystallization work [29] failed to obtain $Cu^{2+}$-containing structures although good crystals were obtained in a $Cu^{2+}$-containing crystallization solution (PDB file 4NYO). In addition, metal analysis of *Ec*CutA confirmed previously that the protein in its apo-state was free of any divalent ions ($Cu^{2+}$, $Ni^{2+}$, $Zn^{2+}$, $Mn^{2+}$, and $Cd^{2+}$) [8].

Our results in both cyanobacteria and *E. coli* ruled out a role for the corresponding PII-like CutA proteins in heavy metal resistance, while only the gene product of *cutA2* was necessary to confer such heavy metal tolerance to *E. coli* cells. In addition, we could not find within the *S. elongatus* genome an orthologous gene for *cutA2* (BLASTp search using *E. coli* CutA2 protein sequence), thus excluding any *cutA2* influence. It is worth noting that the previous report for involvement of PII-like *Ec*CutA in heavy metal tolerance [9] was based on the genetic location of the *E. coli cutA1* gene (PII-like) next to the suppressor copper sensitivity *cutA2* gene, also known as *scsB* and *dsbD* [37–39], but that a rigorous physiological examination of the *cutA1* mutant was lacking. Therefore, the annotation of the PII-like CutA as divalent ion tolerance protein is both misleading and clearly at odds with our results and those previously obtained in other diverse biological systems. In agreement with our phenotypic observations, the level of gene expression of *cutA1* in the plant pathogen *Xylella fastidiosa* showed no significant change in two different media under elevated levels of external $Cu^{2+}$ concentrations [40], while in the higher plant *Arabidopsis thaliana*, the *cutA* knockout lines revealed that *At*CutA is not involved in $Cu^{2+}$ tolerance [41].

While our phenotypic analysis of cyanobacterial *cutA* mutants rule out the involvement of CutA in heavy metal (divalent metal) tolerance, incomplete segregation of the inactivated *cutA* allele in *Nostoc* highlights its biological importance for multicellular lifestyle in this cyanobacterium, similar to *glnB* [42,43] and *sbtB* (ORF *all2133*; unpublished data from K.A. Selim) encoding for canonical PII and PII-like proteins, respectively, in *Nostoc* spp. [42]. Altogether, our biochemical and physiological results emphasize the need to search for the signals and effector molecules regulating CutA activity and the specific cellular processes targeted by this enigmatic protein.

In summary, our present and previous structural studies of various proteins of the PII superfamily [2,3,6,7,31–33,44,45], including canonical PII, SbtB, CPII, PstA, and CutA proteins, indicate that the trimeric assembly is a key evolutionary feature of this superfamily. In fact, this trimeric assembly is also found in cases where the PII domain is a regulatory element controlling the oligomeric state of larger, multidomain proteins such as Nif-3 and HisG [46–48]. These features make the members of the PII superfamily a treasure chest for use in protein design and for the development of *in vivo* and *in vitro* biosensors. While canonical PII proteins have indeed been used to develop *in vivo* 2-OG biosensors [49–51], a CutA protein was engineered to produce a novel tetrahedral two-protein co-assembling nanomaterial, with potential use in targeted drug delivery [52,53]. Further biophysicochemical studies will be necessary for understanding the full breadth of functional roles covered by the PII superfamily in central metabolism and its potential for biotechnological applications.

## Materials and Methods

### Bioinformatic analysis

The analysis was carried out using tools offered within the MPI Bioinformatics Toolkit [19]. The HHpred search to identify distant homologs of CutA was seeded with the *E. coli* CutA protein and was performed against the PDB_mmCIF70 (sequences of protein databank structures clustered at 70% sequence identity) profile HMM database using default settings. To gather CutA homologs, we searched the nr90 sequence database, a version of the nonredundant protein sequence database filtered to a maximum pairwise sequence identity of 90%, using two iterations of PSI-BLAST [18], with an E-value inclusion cutoff of 1e-3. The obtained full-length sequences were clustered in CLANS [21] at a P-value cutoff of 1e-32, with attract value = 2 and repulse value = 20.

### Generation of *cutA* knockout and knockdown strains of cyanobacteria

To inactivate *cutA* in *S. elongatus* PCC 7942, the wild-type strain was transformed with cosmid 8S34-E4 taken from the Unigene set (an existing insertion mutant library for *S. elongatus* [23]) which carries the *cutA* ORF inactivated with a Tn5 transposon conferring resistance to kanamycin. Transformation of *S. elongatus* was achieved using standard protocols [54] and resulted in two classes of transformants: double recombinants that substitute the wild-type allele by the mutant allele, and single recombinants that result from the integration of the entire cosmid with duplication of the targeted gene. The two types of transformants were distinguished by the Cm marker on the cosmid backbone, which is present only in single recombinants. Two clones $Km^R$ $Cm^S$ were isolated, and the correct inactivation of *cutA* was verified by PCR with primers 1F and 1R (Table S2). A second similar alternative for knocking out the *S. elongatus cutA* gene also was successful. In this alternative, the insertion of the kanamycin resistance cassette relied on PCR amplifications using three primer pairs (oligonucleotides used in the present study are shown in Table S2), 1238/1239, 1240/1241, and 1242/1243, on cloning into the pUC19 plasmid using Gibson assembly, and on transformation of *S. elongatus* PCC 7942 with the resulting plasmid, essentially as described for *Synechocystis* sp. PCC 6803 [7]. The mutants were selected on BG11 plates supplemented with 50 $\mu g \cdot mL^{-1}$ kanamycin and verified by PCR with primer pair 1240/1243.

Attempts to knockout the *alr7093* gene (encoding CutA) of *Nostoc* sp. PCC 7120 used the double recombination strategy and kanamycin selection for recombinants. The pRL277 plasmid hosting the upstream and downstream regions of *alr7093* flanking the kanamycin resistance cassette was introduced into the cyanobacteria via conjugation [26]. Nevertheless, full segregation of *alr7093* mutant alleles was not achieved (Fig. 3A).

For *cutA* knockdown in *Nostoc* sp. PCC 7120, a previously reported strategy [25,26] was used in which the cells are transformed with *AS cutA*-pAM1956 plasmid (generated based on pAM1956 plasmid [25]) hosting in reverse-complement orientation *alr7093* (encoding *Ns*CutA) amplified with the 1178/1179 primer pair, placing the reverted gene under the control of a strong copper-inducible promoter ($P_{petE}$), which was cloned using primer pair Fw/1777. The amplified PCR products ($P_{petE}$ and the reverse-complement *alr7093*) were cloned into pAM1956 using Gibson assembly. In the pAM1956-*cutA* construct, a promoter-less *GFPmut2* gene is immediately downstream the antisense-producing *cutA* gene and is co-expressed with *antisense cutA* under $P_{petE}$ control, so that the expression of the downstream green fluorescent protein (GFP) was used as a reporter for production of the *alr7093* antisense-RNA. The *AS cutA* mutant was selected on BG11 plate (without $Cu^{2+}$) supplemented with 50 $\mu g \cdot mL^{-1}$ kanamycin and verified via PCR with primer pair Fw/1179.

## Assessment of metal tolerance of cyanobacterial cells

*S. elongatus* strains were routinely grown photoautotrophically on BG11 medium while shaking at 30°C under continuous illumination conditions ($\sim 30\ \mu E\ m^{-2} \cdot s^{-1}$) [55], checking cell density as $OD^{750}$. Culture biomass was estimated as chlorophyll content from the $OD^{665}$ of a centrifuged methanolic extract of 0.1 mL of culture [56]. In one approach to assess metal tolerance, cultures were adjusted to 10 µg chlorophyll·$mL^{-1}$, and then, serial dilutions of the cell suspension were prepared and equal size drops of these dilutions were seeded on solid 1% agar plates of BG11 medium supplemented with $CuSO_4$, $ZnSO_4$ or $CoCl_2$ as indicated. Plates were photographed after 5 days of growth at 30 °C under constant illumination. In a second approach, exponentially growing wild-type cells of either *S. elongatus* or *Nostoc* sp. and the respective knockout or knockdown strains were subjected to heavy metals stress by supplementing BG11 media with one of the following heavy metals: $Cu^{2+}$, $Fe^{2+}$, $Pb^{2+}$, $Co^{2+}$, $Zn^{2+}$, $Cr^{2+}$, $Cd^{2+}$, $Ni^{2+}$, and $Mn^{2+}$ (used as the chlorides, except $Cu^{2+}$ and $Fe^{2+}$, which were used as the sulfate) at 2.5 – 50 µm concentrations in 24-well plates. The survival of heavy metal-stressed cells was checked by the drop-plating assay [7], where drops from each treatment were spotted on BG11 agar plates in absence of heavy metals and incubated at 28 °C under a light intensity of 30–50 µmol photons $m^{-2} \cdot s^{-1}$ for one week. Growth rates of cells cultivated in liquid medium under stress of 25 µm $Cu^{2+}$ metal were recorded by measuring the increase in optical density at 750 nm ($OD^{750}$) at various time points.

## *Escherichia coli* copper tolerance assays and complementation studies

Wild-type *E. coli* (strain BW25113) and the corresponding *cutA*-lacking strain JW4097 of the Keio collection [27], in which *cutA1* is replaced by a kanamycin resistance cassette, were provided by Dr. F.J. Mojica, Universidad de Alicante. Both strains were lysogenized for phage λDE3 using the λDE3 Lysogenization Kit (as recommended in the kit, from Novagen, Merck Life Sciences, Madrid, Spain), thus to allow T7 promoter-dependent expression of pET15b-carried genes. *E. coli cutA1* and *cutA2* genes (EG12177 and EG12178 Ecogene identification codes, http://ecogene.org/), encoding, respectively, PII-like CutA1 and DsbD, were separately PCR-cloned from genomic DNA of *E. coli* K12 into the *Nde*I-*Bam*HI sites of pET15b, using in the PCR amplification step Deep Vent DNA polymerase (New England Biolabs) and the primer pair *cutA1-Nde*I-F/*cutA1-Bam*HI-R for *cutA1*, and *cutA2-Nde*I-F/*cutA2-Bam*HI-R for *cutA2* (Table S2). The resulting plasmids pET*cutA1* and pET*cutA2* encode the corresponding proteins N-terminally fused to the MGSSHHHHHHSSGLVPRGSH tag. Automated fluorescent Sanger sequencing was used to confirm the correctness of the constructions, the absence of mutations, to determine the insertion point of the kanamycin resistance cassette in the genome of the JW4097 *E. coli* strain, and the lack of insertion of this cassette in the wild-type strain (primers P1 and P2, Table S2, were used for PCR amplification of the region including the insertion point). Standard procedures were used for making electrocompetent cells and for electrotransformation with plasmids, selecting transformants by culturing under aeration at 37°C in LB-ampicillin.

For $Cu^{2+}$ resistance assays, the λDE3-lysogenized BW25113 (WT) and JW4097 cells were grown to identical cell densities (estimated from $OD^{600}$, assuming that 1 OD (for 1-cm light path) corresponds to $10^9$ cells·$mL^{-1}$ [56]), and then, they were diluted to the indicated concentrations in sterile water. Equal size drops of these dilutions were seeded on LB-agar supplemented with or without 3 mm

$CuSO_4$. Growth in individual drops was checked by naked eye after 18 h of plate incubation at 37 °C. Similar assays were carried out with the lysogenized JW4097 strain that had been transformed with either 'empty' parental pET15b or the pET*cutA1* or pET*cutA2* plasmids.

## Production of recombinant CutA proteins

For production of *Se*CutA $His_6$-tagged N-terminally, the *S. elongatus cutA* gene (Cyanobase identifier, *Synpcc7942_2261*), PCR-amplified from genomic DNA using Deep Vent DNA polymerase (New England Biolabs) and the CutA-NdeI-F and CutA-BamHI-R oligonucleotides (Table S2), was cloned into pJET1.2/blunt with the CloneJET PCR Cloning Kit (Thermo Fisher Scientific). Then, after release from this plasmid by *Nde*I/*Bam*HI double digestion, it was subcloned directionally into the corresponding sites of pET28a (Novagen) using Quick Ligase (New England Biolabs). The resulting plasmid encoding *Se*CutA N-terminally fused to the MGSSHHHHHHSSGLVPRGSH tag was used to transform *E. coli* BL21(DE3) cells (Invitrogen), which were grown at 37°C to a cell density of 0.7 $OD^{600}$ in 0.5-L well-aerated cultures in liquid LB-kanamycin (50 μg·mL$^{-1}$) medium. Then, the temperature was lowered to 20°C, 0.1 mM IPTG was added, and the culture was continued at 20°C for 20 h.

Subsequent steps were at 0°C. Centrifugally harvested cells from a 500-mL culture were suspended in 12 mL of 0.1 M Tris/HCl pH 8, 0.15 M NaCl, and 1 mM EDTA and were disrupted by sonication and centrifuged (30 min, 13 500 ***g***), and the supernatant was applied to a 1-mL Ni-chelate HisTrap-HP column fitted in an ÄKTA-FPLC system (both from GE Healthcare), followed by a 30-mL wash with the same buffer with the EDTA replaced by 20 mM imidazole. *Se*CutA was eluted with an 80-mL linear gradient of this buffer containing 20–500 mM imidazole, collecting fractions. The purest fractions (purity monitored by SDS/PAGE) were pooled, the imidazole was removed, and the protein was concentrated to ≥ 7 mg protein·mL$^{-1}$ by centrifugal ultrafiltration (Amicon Ultra, from Millipore; 10 kDa cutoff membrane).

For production of C-terminally Strep-tagged CutA proteins from *Nostoc* sp. PCC 7120 (*Ns*CutA; encoded by ORF *alr7093*) and from *S. elongatus*, the corresponding genes were cloned into pASK-IBA3 as described previously [7,57,58], using, respectively, primer pairs 1414/1415 and 1314/1315 (Table S2). The expression and purification of these strep-tagged CutA proteins were also achieved as described [7,57–59].

## Protein crystallization, X-ray diffraction, and structure determination

Crystallization trials were performed via vapor diffusion at 20–21 °C in 96-well sitting-drop plates hosting drops composed of 0.4 μL of reservoir solution (Table 1) and 0.4 μL of protein solution containing either 7–10 mg·mL$^{-1}$ of *Se*CutA (His-tagged) in 0.1 M Tris/HCl pH 8, 0.15 M NaCl, or 20 mg·mL$^{-1}$ of *Ns*CutA in 25 mM Tris/HCl pH 8.0, 37.5 mM NaCl. *Se*CutA yielded crystals (called here SeL0, SeL1, SeL2, and SeL3, to reflect the number of Bis-Tris molecules bound per trimer, see below) under four conditions (Table 1). For *Ns*CutA, orthorhombic (Ns1) and trigonal (Ns2) crystal forms were obtained with the reservoir solutions listed in Table 1. Diffraction experiments were conducted at 100 K using the synchrotron beamlines and detectors indicated in Table 1.

For *Se*CutA, data were obtained at up to 1.17–2.00 Å resolution and were processed and scaled with XIA2 [60] and AIMLESS [61] (Table 1). Phasing was achieved by molecular replacement with Phaser [62,63] using a single subunit of *E. coli* CutA (PDB 1NAQ) [8] as search model, yielding one CutA homotrimer in the ASU. After refinement (see below), single-subunit models from these crystals were used for phasing in the other crystals of the same protein, yielding single trimers in the ASUs of the other three crystals of *Se*CutA. Rigid body refinement with REFMAC5 [64] (used for optimizing the positions of the three subunits in each trimer and also for model refinement) was followed by alternative iteratively cycles of restrained refinement and of manual model building with Coot [65], with incorporation of ligands when necessary. Isotropic B factors and anisotropic refinement (with Refmac5) were used in the last steps of refinement of SeL2, whereas for SeL0, SeL1, and SeL3, TLS was used in the last refinement steps, the TLS groups being chosen with the TLSMD server (http://skuld.bmsc.washington.edu/˜tlsmd/). The stereochemistry of each model was checked and improved with PDB_REDO [66].

Data obtained for the two *Ns*CutA crystal forms were integrated and scaled using XDS [67] (Table 1). Molecular replacement with MOLREP [68] using monomers of *Thermus thermophilus* CutA (PDB: 1V6H) as a search model located six chains in the ASU of both crystal forms. After initial rigid body refinement with REFMAC5 [64], chain F of crystal Ns2 was poorly defined in the electron density but could be refined using NCS restraints; both structures were completed via cyclic manual modeling with Coot and refinement with REFMAC5.

For all datasets, all diffraction data were used throughout the refinement processes, except the 5% randomly selected data used to calculate the $R_{free}$. Geometry analysis of the protein main chain torsion angles with Rampage [69] and PROCHECK [70] revealed excellent values for the models. Refinement and Ramachandran statistics are shown in Table 1.

## Other methods

Size-exclusion chromatography of His-tagged *Se*CutA (0.02–0.1 mg applied in 0.05 mL) or of standards of known mass was carried out on a Superdex™ 75 10/300 GL column mounted on a ÄKTA FPLC system (column and system from GE Healthcare; Barcelona, Spain) and run with a solution of 0.1 M Tris/HCl pH 8/0.15 M NaCl delivered at a flow rate of 0.4 mL·min$^{-1}$, monitoring continuously the optical absorption of the effluent at 280 nm. A semilogarithmic plot of molecular mass (logarithmic scale) versus elution volume (linear scale) was used to estimate the oligomeric nature of *Se*CutA in solution. Alternatively, when indicated, a Superdex$^{TM}$ 200 PC 3.2/30 column (2.4 mL) was used as previously reported [7,57,58].

Isothermal titration calorimetry was performed using VP-ITC (MicroCal, Malvern Panalytical, Malvern, UK) as described previously [7,57,59], after exhaustive dialysis of CutA protein versus 25 mM Tris/HCl pH 7.8 supplemented with 200 mM NaCl and Chelex 100 to remove any contamination of divalent cations.

Protein was assayed according to Bradford [71], with a commercial reagent from Bio-Rad (Alcobendas, Spain), using bovine serum albumin as a standard. SDS/PAGE was carried out in 15% polyacrylamide gels with Coomassie staining [72]. Amino acid sequence alignments were performed with Clustal omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Graphical representations of protein structures were prepared with PyMOL (http://www.pymol.org/) or UCSF Chimera (http://www.rbvi.ucsf.edu/chimera). Root mean square deviation (rmsd) values were calculated by superimposition of structures using Coot [65].

## Acknowledgments

This work was supported by grants from DFG to K.F. (Fo195/9-2, RTG 1708-2), from DAAD to K.A.S, from the Spanish Government to V.R. and C.M-M. (BFU2014-58229-P, BFU2017-84264-P) and to A.C. (BFU2015-66360-P), and from BioStruct-X (EU) to V.R. for synchrotron access (grant agreement N°283570, proposal 7687), and by Open Access Publishing of Tübingen University. The authors are thankful to Reinhard Albrecht and Nadine Gougeard for technical support, crystallographic sample preparation and assistance with diffraction data collection, and to the IBV-CSIC crystallographic facility for crystal growth. We are grateful to the staff of beamline X10SA/SLS, Diamond (Oxfordshire, UK) and Alba (Barcelona, Spain) synchrotrons for excellent technical support. Open access funding enabled and organized by Projekt DEAL.

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

KAS, LT, KF, and VR conceived and initiated the project. VA performed and wrote the bioinformatic analysis. KAS, JE, and AC generated the different *cutA* knockout mutants of *S. elongatus,*while KAS generated the *cutA* knockdown mutant in *Nostoc*. KAS and LT designed, performed, and interpreted the biochemical, structural, and physiological experiments, with inputs from VR, JE, and AC to LT in the analysis and interpretation of *S. elongatus* experiments performed in Spain. CM-M and MDH supported the structural studies (X-ray crystallography). KAS, MDH, KF, LT, CM-M, and VR analyzed data. KAS, LT, and VR wrote advanced drafts of the manuscript. VR merged the final version of the manuscript, while CM-M and KAS were highly involved with making and editing the figures and tables. All authors but particularly MDH and KF commented and edited on the manuscript, analyzed the results, and approved the final version of the manuscript.

## References

1 Stadtman ER (2001) The story of glutamine synthetase regulation. *J Biol Chem* **276**, 44357–44364.

2 Selim KA, Ermilova E & Forchhammer K (2020) From cyanobacteria to Archaeplastida: new evolutionary insights into PII signalling in the plant kingdom. *New Phytol* **227**, 722–731.

3 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.

4 Thomas G, Coutts G & Merrick M (2000) The *glnKamtB* operon. A conserved gene pair in prokaryotes. *Trends Genet* **16**, 11–14.

5 Kinch LN & Grishin NV (2002) Expanding the nitrogen regulatory protein superfamily: Homology detection at below random sequence identity. *Proteins* **48**, 75–84.

6 Forchhammer K & Lüddecke J (2016) Sensory properties of the PII signalling protein family. *FEBS J* **283**, 425–437.

7 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**, E4861–E4869.

8 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**, 45999–46006.

9 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**, 1127–1137.

10 Navaratnam DS, Fernando FS, Priddle JD, Giles K, Clegg SM, Pappin DJ, Craig I & Smith AD (2000) Hydrophobic protein that copurifies with human brain acetylcholinesterase: amino acid sequence, genomic organization, and chromosomal localization. *J Neurochem* **74**, 2146–2153.
11 Perrier AL, Cousin X, Boschetti N, Haas R, Chatel JM, Bon S, Roberts WL, Pickett SR, Massoulié J, Rosenberry TL & *et al*. (2000) Two distinct proteins are associated with tetrameric acetylcholinesterase on the cell surface. *J Biol Chem* **275**, 34260–34265.
12 Perrier AL, Massoulié J & Krejci E (2002) PRiMA: the membrane anchor of acetylcholinesterase in the brain. *Neuron* **33**, 275–285.
13 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**, 5195–5207.
14 Zhao Y, Wang Y, Hu J, Zhang X & Zhang YW (2012) 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* **287**, 11141–11150.
15 Yang J, Yang H, Yan L, Yang L & Yu L (2009) Characterization of the human CUTA isoform2 present in the stably transfected HeLa cells. *Mol Biol Rep* **36**, 63–69.
16 Naylor MJ, Oakes SR, Gardiner-Garden M, Harris J, Blazek K, Ho TW, Li FC, Wynick D, Walker AM & Ormandy CJ (2005) Transcriptional changes underlying the secretory activation phase of mammary gland development. *Mol Endocrinol* **19**, 1868–1883.
17 Laskowski RA, Watson JD & Thornton JM (2003) From protein structure to biochemical function? *J Struct Funct Genomics* **4**, 167–177.
18 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K & Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421.
19 Zimmermann L, Stephens A, Nam SZ, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN & Alva VA (2018) Completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* **430**, 2237–2243.
20 Fox NK, Brenner SE & Chandonia JM (2014) SCOPe: Structural Classification of Proteins–extended, integrating SCOP and ASTRAL data and classification of new structures. *Nucleic Acids Res* **42**, D304-D-309.
21 Frickey T & Lupas A (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. *Bioinformatics* **20**, 3702–3704.
22 Fournier GP & Poole AM (2018) A briefly argued case that Asgard archaea are part of the eukaryote tree. *Front Microbiol* **9**, 1896.
23 Holtman CK, Chen Y, Sandoval P, Gonzales A, Nalty MS, Thomas TL, Youderian P & Golden SS (2005) High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. *DNA Res* **12**, 103–115.
24 Rubin BE, Wetmore KM, Price MN, Diamond S, Shultzaberger RK, Lowe LC, Curtin G, Arkin AP, Deutschbauer A & Golden SS (2015) The essential gene set of a photosynthetic organism. *Proc Natl Acad Sci USA* **112**, E6634–E6643.
25 Srivastava A, Brilisauer K, Rai AK, Ballal A, Forchhammer K & Tripathi AK (2017) Down-regulation of the alternative sigma factor SigJ confers a photoprotective phenotype to *Anabaena* PCC 7120. *Plant Cell Physiol* **58**, 287–297.
26 Bornikoel J, Staiger J, Madlung J, Forchhammer K & Maldener I (2018) LytM factor Alr3353 affects filament morphology and cell-cell communication in the multicellular cyanobacterium *Anabaena* sp. PCC 7120. *Mol Microbiol* **108**, 187–203.
27 Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL & Mori H (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2** (2006), 0008.
28 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**, 8960–8972.
29 Bagautdinov B (2014) The structures of the CutA1 proteins from *Thermus thermophilus* and *Pyrococcus horikoshii*: characterization of metal-binding sites and metal-induced assembly. *Acta Crystallogr F Struct Biol Commun* **70**, 404–413.
30 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**, 322–336.
31 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**, 4013–4030.
32 Forchhammer K & Selim KA (2020) Carbon/nitrogen homeostasis control in cyanobacteria. *FEMS Microbiol Rev* **44**, 33–53.
33 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**, 2888–2901.
34 Gundlach J, Dickmanns A, Schröder-Tittmann K, Neumann P, Kaesler J, Kampf J, Herzberg C, Hammer

E, Schwede F, Kaever V *et al.* (2015) Identification, characterization, and structure analysis of the cyclic di-AMP-binding PII-like signal transduction protein DarA. *J Biol Chem* **290**, 3069–3080.
35 Müller M, Hopfner KP & Witte G (2015) c-di-AMP recognition by *Staphylococcus aureus* PstA. *FEBS Lett* **589**, 45–51.
36 Tanaka Y, Tsumoto K, Nakanishi T, Yasutake Y, Sakai N, Yao M, Tanaka I & Kumagai I (2004) Structural implications for heavy metal-induced reversible assembly and aggregation of a protein: the case of *Pyrococcus horikoshii* CutA. *FEBS Lett* **556**, 167–174.
37 Gupta SD, Wu HC & Rick PD (1997) A *Salmonella typhimurium* genetic locus which confers copper tolerance on copper-sensitive mutants of *Escherichia coli*. *J Bacteriol* **179**, 4977–4984.
38 Stirnimann CU, Rozhkova A, Grauschopf U, Grütter MG, Glockshuber R & Capitani G (2005) Structural basis and kinetics of DsbD-dependent cytochrome c maturation. *Structure* **13**, 985–993.
39 Furlong EJ, Choudhury HG, Kurth F, Duff AP, Whitten AE & Martin JL (2018) Disulfide isomerase activity of the dynamic, trimeric *Proteus mirabilis* ScsC protein is primed by the tandem immunoglobulin-fold domain of ScsB. *J Biol Chem* **293**, 5793–5805.
40 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**, 1145–1157.
41 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**, 856–867.
42 Hanson TE, Forchhammer K, Tandeau de Marsac N & Meeks JC (1998) Characterization of the *glnB* gene product of *Nostoc punctiforme* strain ATCC 29133: *glnB* or the PII protein may be essential. *Microbiology* **144**, 1537–1547.
43 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**, 33641–33648.
44 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 USA* **104**, 17644–17649.
45 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 USA* **107**, 15397–15402.
46 Godsey MH, Minasov G, Shuvalova L, Brunzelle JS, Vorontsov II, Collart FR & Anderson WF (2007) The 2.2 Å resolution crystal structure of *Bacillus cereus* Nif3-family protein YqfO reveals a conserved dimetal-binding motif and a regulatory domain. *Protein Sci* **16**, 1285–93.
47 Chellamuthu VR, Alva V & Forchhammer K (2013) From cyanobacteria to plants: conservation of PII functions during plastid evolution. *Planta* **237**, 451–462.
48 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**, 2789–2793.
49 Lüddecke J & Forchhammer K (2013) From PII signaling to metabolite sensing: a novel 2-oxoglutarate sensor that details PII-NAGK complex formation. *PLoS One* **8**, e83181.
50 Lüddecke J, Francois L, Spät P, Watzer B, Chilczuk T, Poschet G, Hell R, Radlwimmer B & Forchhammer K (2017) PII protein-derived FRET sensors for quantification and live-cell imaging of 2-oxoglutarate. *Sci Rep* **7**, 1437.
51 Chen HL, Latifi A, Zhang CC & Bernard CS (2018) Biosensors-based in vivo quantification of 2-oxoglutarate in cyanobacteria and proteobacteria. *Life (Basel)* **8**, E51.
52 King NP, Bale JB, Sheffler W, McNamara DE, Gonen S, Gonen T, Yeates TO & Baker D (2014) Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108.
53 Bale JB, Park RU, Liu Y, Gonen S, Gonen T, Cascio D, King NP, Yeates TO & Baker D (2015) Structure of a designed tetrahedral protein assembly variant engineered to have improved soluble expression. *Protein Sci* **24**, 1695–1701.
54 Clerico EM, Ditty JL & Golden SS (2007) Specialized techniques for site-directed mutagenesis in cyanobacteria. *Methods Mol Biol* **362**, 155–171.
55 Labella JI, Cantos R, Espinosa J, Forcada-Nadal A, Rubio V & Contreras A (2017) PipY, a member of the conserved COG0325 family of PLP-binding proteins, expands the cyanobacterial nitrogen regulatory network. *Front Microbiol* **8**, 1244.
56 Elbing K & Brent R. (2002) Growth in liquid media. In Current Protocols in Molecular Biology Chapter 1, Unit 1.2.
57 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**, 18985.
58 Selim KA, Lapina T, Forchhammer K & Ermilova E (2020) 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* **287**, 465–482.

59 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**, 790.
60 Winter G, Lobley CMC & Prince SM (2013) Decision making in xia2. *Acta Cryst Sect D, Biol Cryst* **69**, 1260–1273.
61 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Cryst Sect D, Biol Cryst* **69**, 1204–1214.
62 McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Cryst Sect D, Biol Cryst* **63**, 32–41.
63 McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC & Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674.
64 Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F & Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Cryst Sect D, Biol Cryst* **67**, 355–367.
65 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Cryst Sect D, Biol Cryst* **66**, 486–501.
66 Joosten RP, Salzemann J, Bloch V, Stockinger H, Berglund A-C, Blanchet C, Bongcam-Rudloff E, Combet C, Da Costa AL, Deleage G *et al.* (2009) PDB_REDO: automated re-refinement of X-ray structure models in the PDB. *J Appl Crystallogr* **42**, 376–384.
67 Kabsch W (2010) XDS. *Acta Cryst Sect D, Biol Cryst* **66**, 125–132.
68 Vagin A & Teplyakov A (2010) Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr* **66**, 22–25.
69 Lovell SC, Davis IW, Arendall WB, de Bakker PIW, Word JM, Prisant MG, Richardson JS & Richardson DC (2003) Structure validation by $C_{alpha}$ geometry: phi, psi and $_{Cbet}$a deviation. *Proteins* **50**, 437–450.
70 Laskowski RA, MacArthur MW, Moss DS & Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* **26**, 283–291.
71 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
72 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
73 Labella JI, Llop A & Contreras A (2020) The default Cyanobacterial Linked Genome (dCLG): an interactive platform based on cyanobacterial linkage networks to assist functional genomics. *FEBS Lett*. **594**,1661–1674.
74 Vijayan V, Jain IH & O'Shea EK (2011) A high resolution map of a cyanobacterial transcriptome. *Genome Biol* **12**, R47.
75 Tan X, Hou S, Song K, Georg J, Klahn S, Lu X & Hess WR (2018) The primary transcriptome of the fast-growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Biotechnol Biofuels* **11**, 218.
76 Jurrus E, Engel D, Star K, Monson K, Brandi J, Felberg LE, Brookes DH, Wilson L, Chen J, Liles K *et al.* (2018) Improvements to the APBS biomolecular solvation software suite. *Protein Sci* **27**, 112–128.
77 Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T & Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* **44**, W344–W350.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1.** Previously determined CutA structures.

**Table S2.** Oligonucleotides used in the present study.

**Fig. S1.** Structural superpositions of *Ns*CutA with *Ec*CutA, with canonical PII and with the SbtB non-canonical member of the PII superfamily.

**Fig. S1.** Alignment of amino acid sequences of CutA proteins from different organisms.