

**Discovery of cyanophycin dipeptide hydrolase enzymes suggests  
widespread utility of the natural biopolymer cyanophycin**

Itai Sharon<sup>a,1</sup>, Geoffrey A. McKay<sup>b,1</sup>, Dao Nguyen<sup>b,2</sup>, T. Martin Schmeing<sup>a,1</sup>

<sup>a</sup>Department of Biochemistry and Centre de recherche en biologie structurale, McGill University, Montréal, QC, Canada, H3G 0B1.

<sup>b</sup>Meakins-Christie Laboratories, Research Institute of the McGill University Health Centre, 1001 Décarie Boulevard, Montréal, Québec, Canada H4A 3J1

<sup>1</sup>I.S. and G.M. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Emails: martin.schmeing@mcgill.ca;  
dao.nguyen@mcgill.ca

**Author Contributions:** I.S. and T.M.S. designed the biochemical study and performed the bioinformatic analysis. I.S. performed all biochemical and structural experiments and data processing. G.M. and D.N. designed all experiments with *Pseudomonas aeruginosa*, which G.M. performed. All authors wrote and edited the manuscript.

**Competing Interest Statement:** The authors declare no competing interests.

**Classification:** Biological sciences; Biochemistry.

**Keywords:** Cyanophycin; dipeptide hydrolase; enzyme discovery; structural biology; *Pseudomonas aeruginosa*.

## Abstract

Cyanophycin is a bacterial polymer mainly used for nitrogen storage. It is composed of a peptide backbone of L-aspartate residues with L-arginines attached to their side chains through isopeptide bonds. Cyanophycin is degraded in two steps: Cyanophycinase cleaves the polymer into  $\beta$ -Asp-Arg dipeptides, which are hydrolyzed into free Asp and Arg by enzymes possessing isoaspartyl dipeptidase activity. Isoaspartyl dipeptidase (IadA) and isoaspartyl aminopeptidase (IaaA) have been shown to degrade  $\beta$ -Asp-Arg dipeptides, but bacteria which encode cyanophycin-metabolizing genes can lack *iaaA* and *iadA* genes. In this study, we investigate a previously uncharacterized enzyme whose gene can cluster with cyanophycin-metabolizing genes. This enzyme, which we name cyanophycin dipeptide hydrolase (CphZ), is specific for dipeptides derived from cyanophycin degradation. Accordingly, a co-complex structure of CphZ and  $\beta$ -Asp-Arg shows that CphZ, unlike IadA or IaaA, recognizes all portions of its  $\beta$ -Asp-Arg substrate. Bioinformatic analyses showed that CphZ is found in very many proteobacteria and is homologous to an uncharacterized protein encoded in the “arginine/ornithine transport” (*aot*) operon of many *Pseudomonas* species, including *Pseudomonas aeruginosa*. *In vitro* assays show that AotO is indeed a cyanophycin dipeptide hydrolase, and *in cellulo* growth experiments show that this enzyme and the *aot* operon allow *P. aeruginosa* to take up and use  $\beta$ -Asp-Arg as a sole carbon and nitrogen source. Together the results establish the novel, highly specific enzyme subfamily of cyanophycin dipeptide hydrolases, suggesting that cyanophycin is potentially used by a much wider range of bacteria than previously appreciated.

**Significance statement:** Cyanophycin is best known for its roles in cyanobacteria, which have developed dedicated pathways around the biopolymer. We have shown that a novel, common, dedicated cyanophycin dipeptide hydrolase family is found in bacteria which possess other cyanophycin metabolizing genes, as well as in those that lack them. The specificity of cyanophycin dipeptide hydrolases is derived from conserved interactions of the active site with both residues of its  $\beta$ -Asp-Arg dipeptide substrate. This specificity is maintained in cyanophycin dipeptide hydrolase enzymes from both cyanophycin producers and scavengers with medical relevance. The prevalence of cyanophycin metabolism genes and the discovery of this dedicated enzyme family highlight the utility of this interesting natural biopolymer.

## **Introduction**

Cyanophycin is a biopolymer produced by a wide range of bacteria found in environments such as marine habitats, sediment and gut microbiota (1-9). It was first discovered around 140 years ago as large dark granules in cyanobacterial cells, visible under a simple light microscope (10). Cyanophycin chains are composed of a backbone of L-aspartate residues with L-arginines attached to each aspartate side chain through isopeptide bonds (i.e.,  $(\beta\text{-Asp-Arg})_n$ ) (11) (Fig. 1). Although most often described as a nitrogen storage polymer (12), cyanophycin can also be useful as a store of carbon and energy (13-15). Cyanophycin as a dynamic nitrogen reservoir is beneficial for cells in a variety of conditions (16-18). For example, nitrogen fixing cyanobacteria can use it to separate (aerobic) photosynthesis from the strictly anaerobic process of nitrogen fixation. By producing cyanophycin, cells can stockpile excess fixed nitrogen during anaerobic periods of a day/night cycle (19) or in anaerobic cell types (3, 20). This fixed nitrogen can then be mobilized and used in aerobic conditions, because cyanophycin catabolism is insensitive to  $\text{O}_2$ .

Cyanophycin is commonly synthesized by cyanophycin synthetase 1 (CphA1) (21). CphA1 has two synthetic active sites which alternately add Asp and Arg to the nascent polymer in ATP-dependent manners (1, 21). Many CphA1s also harbor a hydrolytic active site that generates cyanophycin primers to facilitate synthesis (22). Some lysine can be incorporated into cyanophycin in place of Arg (23), the amount of which depends on the CphA1 enzyme, the native or heterologous host and the growth conditions (24), but it is typically much lower than Arg. An enzyme related to CphA1 (CphA2), found only in cyanobacteria possessing CphA1, can re-polymerize  $\beta\text{-Asp-Arg}$  dipeptides into cyanophycin using its single active site (25, 26).

Bacteria must degrade cyanophycin polymer into its constituent amino acids to access the stored carbon, nitrogen and energy (3, 19). This degradation happens in two steps (Fig. 1). First, the polymer is hydrolyzed into  $\beta\text{-Asp-Arg}$  dipeptides by cyanophycinase (27, 28). Then, hydrolysis of isopeptide bond splits  $\beta\text{-Asp-Arg}$  into Asp and Arg, which can enter primary metabolism (29). Unrelated enzymes capable of catalyzing this second reaction can be grouped under the label “isoaspartyl dipeptidases” (Fig. 1).

Isoaspartyl dipeptidases are common because they can also participate in a fairly widespread protein damage pathway: Proteins become spontaneously damaged by intramolecular rearrangement of Asp or Asn residues, in which the peptide backbone from the Asp or Asn main chain is transferred to the side chain. This lesion can be repaired by L-isoaspartyl O-

methyltransferases (30), or the damaged protein can be degraded by proteases that hydrolyze the protein backbone. However, proteases cannot digest isoaspartyl dipeptides (31, 32), so isoaspartyl dipeptidases (33) must hydrolyze them to prevent accumulation to toxic levels (34). Any amino acid can become attached to the Asp side chain during protein damage, and accordingly isoaspartyl dipeptidases accept a wide range of isoaspartyl dipeptides as substrates (29, 35). Cyanophycin dipeptide degradation is thought to be performed by isoaspartyl dipeptidases moonlighting from their role in the damaged protein pathway.

Two unrelated types of bacterial isoaspartyl dipeptidase enzymes are known: “isoaspartyl dipeptidase” (IadA) (33, 36) and “isoaspartyl aminopeptidase” (IaaA) (29, 37-39). Both can degrade  $\beta$ -aspartyl dipeptides derived from damaged protein, as well as  $\beta$ -Asp-Arg and  $\beta$ -Asp-Lys derived from cyanophycin (40). However, over a quarter of cyanophycin-producing bacteria do not possess genes that encode IadA or IaaA (2, 40). We recently observed that other putative hydrolases can cluster with cyanophycin metabolizing genes in some bacteria that lack IadA and IaaA (40).

In this study, we perform structural, biochemical and bioinformatic studies to interrogate the activity and role of a previously uncharacterized enzyme from *Acinetobacter baylyi* (*AbCphZ*). We show that, unlike the unrelated, characterized isoaspartyl dipeptidases IaaA and IadA, *AbCphZ* is specific for  $\beta$ -Asp-Arg/Lys. Furthermore, we characterize *Pseudomonas aeruginosa* AotO, a previously uncharacterized enzyme commonly found in non-cyanophycin producing proteobacteria, to likewise be a cyanophycin dipeptide hydrolase. We find that genes in the *aot* operon, including *aotO*, allow *P. aeruginosa* to use  $\beta$ -Asp-Arg as a nitrogen and carbon source. Together with previous results showing existence of cyanophycin in disparate environments (8, 9), the results imply that cyanophycin is a common material, and that bacteria may specifically scavenge its  $\beta$ -Asp-Arg degradation product.

## **Results**

### ***AbCphZ* is a $\beta$ -Asp-Arg/Lys dipeptidase**

We recently performed a bioinformatic analyses of co-occurrence and clustering of cyanophycin metabolizing genes (40) and noted that 27% of genomes which encode cyanophycin synthetase lack both *iadA* and *iaaA*. Some cyanophycin synthetase-cyanophycinase clusters instead include a gene annotated as “M14 family metallopeptidase” or “succinylglutamate

desuccinylase/aspartoacylase family protein”. Notably, they include *Acinetobacter baylyi* DSM587, one of the five bacteria noted by Füsler and Steinbüchel (2) to encode CphA1 and cyanophycinase but not IadA nor IaaA. We hypothesized that *Acinetobacter baylyi* DSM587 “M14 family metallopeptidase” (WP\_004925890.1), which shares ~22% identity with *E. coli* succinylglutamate desuccinylase (AstE) (41), may perform hydrolysis of  $\beta$ -Asp-Arg/Lys derived from cyanophycin, and named it *A. baylyi* cyanophycin dipeptide hydrolase CphZ (*AbCphZ*).

We heterologously expressed *AbCphZ* in *E. coli* for activity and structural studies. *AbCphZ* could be purified to homogeneity and migrates as a dimer in size exclusion chromatography (SI Appendix, Fig. S1A). We tested the enzyme’s ability to hydrolyze various  $\beta$ -aspartyl dipeptides. *AbCphZ* displays a clear preference towards  $\beta$ -Asp-Arg/Lys dipeptides, hydrolyzing these robustly, but not other  $\beta$ -aspartyl dipeptides (Fig. 2A). *AbCphZ* also displayed specificity toward the  $\beta$ -linkage, as  $\alpha$ -Asp-Arg was not efficiently hydrolyzed (Fig. 2A). It is also specific towards the Asp residue of the substrate dipeptide, as it has very low activity on N(2)-acetyl-Arg and no detectable activity on N(2)-succinyl-Arg, an intermediate in the AST arginine catabolism pathway, which is very similar in structure to  $\beta$ -Asp-Arg (Fig. 2B) (41). Likewise, it has nearly no activity on the most similar glutamate containing isodipeptide,  $\gamma$ -Glu-Arg. The overall substrate specificity is remarkable, as we are unaware of any other enzyme that has selectivity for  $\beta$ -Asp-Arg/Lys dipeptides. Furthermore, cyanophycin degradation is the only abundant source of  $\beta$ -Asp-Arg/Lys dipeptides we know.

To determine the structural basis for this substrate specificity, we determined the crystal structure of *AbCphZ* at 2.7 Å resolution (Fig. 2C, SI Appendix, Table S1). *AbCphZ* is a homodimer of protomers which contains a large domain (a 9-strand  $\beta$  sheet surrounded by  $\alpha$  helices) and a small domain (made of  $\beta$  strands), the same fold possessed by *E. coli* AstE (PDB code 1YW6, RMSD 3.8 Å across 109 C $_{\alpha}$  pairs, SI Appendix, Fig. S1B) and bovine pancreatic carboxypeptidase A (42, 43) (PDB code 1HEE, RMSD 4.1 Å across 159 C $_{\alpha}$  pairs, SI Appendix, Fig. S1C). As with AstE and carboxypeptidase A, the active site of *AbCphZ* contains a single metal ion liganded by a conserved H-H-E triad (*AbCphZ* H50, E53, and H179, Fig. 2D, SI Appendix, Fig. S1C,D). The density map indicates the presence of another metal ion near Y62 of one protomer and H315 and S330 of another protomer within the CphZ dimer. Inductively coupled plasma mass spectrometry (ICP-MS) showed our CphZ samples contain both zinc and manganese (SI Appendix, Table S2). We verified that the ion in the active site is Zn<sup>2+</sup> using single-wavelength anomalous diffraction

(SI Appendix, Fig. S1E), and modelled  $Mn^{2+}$  in the second metal site (Supplemental Fig. 1F). The conserved E251 is positioned close to the  $Zn^{2+}$  ion and the putative substrate binding site (SI Appendix, Fig. S1D), and is essential for activity (Fig. 2D), like analogous glutamates in other zinc metallopeptidases (44).

To understand how CphZ specifically binds and cleaves cyanophycin-derived dipeptides, we determined the co-complex structure of the catalytically inactive mutant CphZ<sub>E251A</sub> bound to  $\beta$ -Asp-Arg at 2.2 Å resolution (Fig. 2E,F, SI Appendix, Table S1). The maps show no substantial rearrangement of the protein from the unliganded conformation, but clear density for  $\beta$ -Asp-Arg in the active site (Fig. 2E). The carbonyl of the dipeptide's scissile bond is positioned 2.0 Å away from the  $Zn^{2+}$  ion (Fig. 2F), in good position for catalysis (45). The  $Zn^{2+}$  will serve as a Lewis acid and draw electrons from the bond and facilitate its cleavage, presumably by a water molecule activated by E251 (44). The structure also explains how CphZ specifically recognizes and binds the Arg/Lys portion of the substrate: Conserved residues D222 and E223 form a negatively charged pocket suitable for the binding of a positively charged guanidino or amino group (Fig. 2F). In addition, the conserved R102, N112 and R113 bind the Arg carboxyl of the dipeptide in a similar manner to that observed in carboxypeptidase A (SI Appendix, Fig. S1C), orienting the scissile bond for catalysis. Mutation of any of N112, R113, D222 and E223 abolishes enzyme activity, while mutation of S216, which is not as highly conserved, decreases activity to approximately half (Fig. 2D). The Asp residue of the dipeptide interacts with a pocket that shows lower sequence conservation (SI Appendix, Fig. S1G), but D181 and K366 are well conserved and within hydrogen bonding distance from the Asp (Fig. 2F). Isothermal titration calorimetry (ITC) binding assays with  $\beta$ -Asp-Arg, Asn and Arg shows CphZ<sub>E251A</sub> has good affinity for the substrate  $\beta$ -Asp-Arg ( $K_D = 7.6 \pm 0.1 \mu M$ ), and much lower affinity for Arg ( $K_D = 470.8 \pm 16.8 \mu M$ ) and Asn (no binding detected) (SI appendix, Fig. S1H).

### ***P. aeruginosa* AotO is also a cyanophycin dipeptide hydrolase**

The Conserved Domain Database (46) identifies residues 18-280 of AbCphZ as a “M14\_PaAOTO\_like” (accession cd06250) domain. This domain is also present in the uncharacterized protein AotO (47) of *Pseudomonas aeruginosa* (WP\_128550578.1; PA0891 (48); here PaAotO). PaAotO and AbCphZ share 38% sequence identity, with particularly high similarity around the active site (SI Appendix, Fig. S2A). *Pseudomonas* species do not possess genes for

cyanophycin synthesis, but the homology between CphZ and AotO led us to ask whether these enzymes possess similar activity.

AotO is encoded in the *aot* operon, which also encodes a multi-component ABC transporter (AotJQMP) and an Arg-dependent transcription activator (ArgR) (47). AotJQMP is important in *P. aeruginosa* for Arg uptake and its use as a carbon source (47, 49). *P. aeruginosa aotO* does not affect Arg uptake, and its role is unknown (47). To determine whether CphZ and AotO have the same activity, we expressed and purified *PaAotO* and tested its activity towards  $\beta$ -Asp-Arg, N(2)-succinyl arginine and N(2)-acetyl arginine. Like *AbCphZ*, AotO is specific for  $\beta$ -Asp-Arg and hydrolyzes it at a rate comparable to *AbCphZ* and the non-specific isoaspartyl dipeptidases *RhIaaA* and *LmIadA* (40) (Fig. 3A,B).

### **AotO allows *P. aeruginosa* to grow on $\beta$ -Asp-Arg as sole carbon and nitrogen source**

To demonstrate that *P. aeruginosa* can catabolize  $\beta$ -Asp-Arg *in cellulo*, we cultured the wild type PAO1 strain (50) in minimal media supplemented with  $\beta$ -Asp-Arg.  $\beta$ -Asp-Arg alone is sufficient to support bacterial growth, although with a slower doubling time (36.6 h) than with D-glucose and  $\text{NH}_4\text{Cl}$  (7.0 h) (Fig. 3C). Bacterial replication in media containing  $\beta$ -Asp-Arg was confirmed by a marked increase in viable cell counts from  $1.9 \times 10^7$  CFU/ml to  $5.7 \times 10^8$  CFU/ml in 48 hours of flask growth (Fig. 3D). Thus,  $\beta$ -Asp-Arg can be catabolized by *P. aeruginosa* and sustains bacterial growth as the sole carbon and nitrogen source in flasks, as well as in microplates (SI appendix, Fig. S2B,C). The related dipeptide  $\beta$ -Asp-Ala does not support growth alone or with D-glucose, consistent with the substrate specificity of cyanophycin dipeptide hydrolases (SI Appendix, Fig. S2D).

$\beta$ -Asp-Arg supports bacterial growth well in combination with D-glucose, with a clear  $\beta$ -Asp-Arg concentration dependence (Fig. 3E). Doubling times of 17.4 h, 7.3 h and 5.2 h were observed with 0.1 mM, 1 mM and 10 mM  $\beta$ -Asp-Arg respectively. Notably, growth in 20 mM D-glucose and 10 mM  $\beta$ -Asp-Arg is comparable to that in D-glucose and 10 mM  $\text{NH}_4\text{Cl}$ , indicating that  $\beta$ -Asp-Arg serves as a very efficient nitrogen source (Fig. 3C,E).

To confirm that the cyanophycin dipeptide hydrolase AotO is responsible for  $\beta$ -Asp-Arg catabolism in *P. aeruginosa*, we assayed growth of the *aotO* mutant from a sequence-verified transposon mutant library (51). The *aotO* mutant was unable to grow on  $\beta$ -Asp-Arg as sole carbon and nitrogen source (Fig. 3F). *aotM* and *aotP* mutants, which lack functional AotJQMP

transporter, were likewise unable to grow on  $\beta$ -Asp-Arg (Fig. 3F, SI Appendix, Fig. S2E). *aot* transposon mutants grew identically to their wildtype parental strain in media containing  $\text{NH}_4\text{Cl}$  and D-glucose (SI Appendix, Fig. S2F-L), confirming that AotO and the AotJQMP transporter are specifically required for  $\beta$ -Asp-Arg catabolism.

### **The genetic context of *cphZ***

CphZ is a common enzyme: There are nearly 10,000 entries in the Max Planck Institute Bioinformatics Toolkit (52) non-redundant (nr) database with >70% coverage and >30% identity to *AbCphZ*, and over 2,000 in the nr 90% (nr90) database. To better understand the role of CphZ, we performed an analysis of the distribution and genomic localization of *cphZ* relative to other cyanophycin metabolism genes. We searched the RefSeq database (53) for complete bacterial genomes that include *cphZ* and constructed a phylogenetic tree of these using phyloT (54) according to the genome taxonomy database classification (55) (Fig. 3G). This shows CphZ to be principally a proteobacterial enzyme, which is particularly common in the family *Pseudomonadaceae* and in the order *Burkholderiales*. A WebLogo (56) shows high conservation of active site and substrate binding residues (Fig. 3H) in these CphZ sequences. While many *Burkholderiales* encode CphA1 (40), a BLAST (57) search shows that cyanophycin producers are extremely rare in *Pseudomonadaceae*, with only one sequenced strain possessing a *cphA1* gene.

We then constructed a database of all the complete bacterial genomes in the RefSeq database and used Cblaster (58) to analyze the co-occurrence and clustering of *cphA1*, cyanophycinase (*cphB/E/I*) and cyanophycin dipeptide hydrolase (*cphZ/aotO*). Of the 27,349 genomes in that database, 3,095 (~11%) genomes have at least one of *cphA1*, cyanophycinase or *cphZ* (Table 1). *cphA1* is present in 1614 (6%), a cyanophycinase gene is present in 840 (~3%) and *cphZ* is present in 1364 (~5%). However, only 41 genomes encode all three of cyanophycin synthetase, cyanophycinase and cyanophycin dipeptide hydrolase, and all three genes cluster to within 5 kilobases in only 12 of these 41 genomes. Co-occurrence and clustering analyses of cyanophycinase and CphZ show 50 genomes encode cyanophycinase and CphZ (but not CphA1), and that clustering is more common, with 32 of 50 genomes (Table 1) possessing a two-gene cyanophycin-degradation cluster.

### **Discussion**

The experimental and bioinformatic results demonstrate that cyanophycin dipeptide hydrolase is a novel enzyme subfamily that is common and highly specific for  $\beta$ -Asp-Arg/Lys dipeptides. These CphZs are the first known enzymes dedicated to degradation of cyanophycin dipeptides, unlike IaaAs and IadAs, which degrade cyanophycin metabolites as a secondary function. *PaAotO* can equally be named *PaCphZ* since it also a founding member of the cyanophycin dipeptide hydrolase subfamily. Two functionally uncharacterized proteins with structures deposited in the PDB may also be CphZs: *Shewanella amazonensis* “succinylglutamate desuccinylase aspartoacylase” (accession code 3FMC) shares 27% identity with *AbCphZ*, and the structures have a root mean squared deviation (rmsd) of 1.8 Å across 220 C $_{\alpha}$  pairs. *Shewanella frigidimarina* “putative succinylglutamate desuccinylase/aspartoacylase” (3LWU) shares 29% identity and has a 1.7 Å rmsd. These two gammaproteobacterial species also encode cyanophycinases. Both CphZ homologs share lower identity with *AbCphZ* or *PaAotO* than the cutoff we use in our bioinformatic analysis, but their identities to *AbCphZ* are not very different from that of *PaAotO* (38%) which we confirmed is a cyanophycin dipeptide hydrolase. Inspection of the structures 3FMC and 3LWU reveals conservation of the residues that form the Arg and Asp binding pockets, suggesting these enzymes are likely CphZs and that there may be even more CphZs in nature than our analyses indicate.

Cyanophycin dipeptide hydrolases are common in proteobacteria, which represent over half the bacterial sequences in genomic databases. Like IaaA and IadA, CphZ is mostly found in strains that cannot produce cyanophycin (Table 1; (40)). A study identifying core genes among *Pseudomonas* groups found *aotO* to be conserved among common species, and that the whole *aot* operon also shows good conservation (59). Cyanophycin-degrading *Proteobacteria* have been isolated from a variety of environments and microbiomes (60-64), cyanophycin itself has been found in disparate environments (8, 9), and many bacteria encode a secreted version of cyanophycinase, CphE (Table 1, (60)). The observation that *cphZ* is more commonly found in genomes that lack a recognizable cyanophycinase gene (1314 genomes with *cphZ* but no cyanophycinase gene; 50 with *cphZ* and a cyanophycinase gene) suggests these bacteria (including *Pseudomonas*) live in communities with other species that degrade polymer into dipeptide (63). There is no known cyanophycin export process, so presumably cyanophycin in the environment comes from cell lysis of producer bacteria. Extracellular cyanophycinase degrades this

environmental cyanophycin to  $\beta$ -Asp-Arg, which is imported and used by CphZ (or IaaA / IadA) encoding species.

The frequency of cyanophycin metabolizing genes (Table 1) indicates that it is clearly advantageous and common for bacteria to have mechanisms capable of scavenging cyanophycin and its breakdown product  $\beta$ -Asp-Arg. Indeed, the analogous bioinformatic analyses of glycogen metabolizing genes shows cyanophycin metabolizing genes are present at a ~1:3 ratio to those of glycogen, which is considered a ubiquitous biopolymer (SI Appendix, Table S3).

Herein, we characterized growth on  $\beta$ -Asp-Arg of one species that possesses a cyanophycin dipeptide hydrolase but not a recognizable cyanophycinase:  $\beta$ -Asp-Arg can be used by the human pathogen *P. aeruginosa* as the sole carbon source and as an excellent sole nitrogen source, nearly equivalent to  $\text{NH}_4\text{Cl}$ . Growth on  $\beta$ -Asp-Arg appeared biphasic (Fig. 3E), suggesting a possible induction of processes required for rapid metabolism of  $\beta$ -Asp-Arg, perhaps *aot* operon upregulation by the ArgR regulator (47). Transposon mutation of *aotO* or of the L-Arg transporter subunits *aotM* and *aotP* abolishes the ability to use  $\beta$ -Asp-Arg as a nitrogen and/or carbon source. These data confirm the *in vivo* cyanophycin dipeptide hydrolase function of *aotO* and also imply that the transporter AotJQMP can import  $\beta$ -Asp-Arg, not only Arg. Multiple mechanisms for L-Arg uptake have been identified, including by *P. aeruginosa* ArcD, an arginine/ornithine antiporter (65). ArcD seems unable to import  $\beta$ -Asp-Arg, because the *aotM* mutant cannot grow on  $\beta$ -Asp-Arg, but can grow on Arg. Given the difference in size and nature between Arg and  $\beta$ -Asp-Arg, it is remarkable that AotJQMP can import both compounds. Furthermore, given the apparent advantage strains can derive from use of environmental cyanophycin dipeptide, inhibition of its uptake or hydrolysis by pathogenic strains could be a promising strategy for limiting their proliferation within complex microbiota.

## **Conclusion**

In conclusion, although cyanophycin is best known for its roles in cyanobacteria, the discovery of a common, dedicated cyanophycin dipeptide hydrolase family in both cyanophycin producers and scavengers with medical relevance builds on the recent reports of wide environmental availability to highlight the utility of this interesting natural biopolymer.

## **Materials and methods**

### Cloning, protein expression and purification

The cloning, and protein expression and purification performed in this study were like that described previously for other cyanophycin metabolizing enzymes (1, 22, 26, 28). The gene encoding *AbCphZ* (WP\_004925890.1) was amplified from *Acinetobacter baylyi* DSM587 genomic DNA (DSMZ, Germany). The codon-optimized synthesis of the gene encoding *PaAotO* (WP\_128550578.1) was commissioned from Bio Basic, Canada. The genes encoding *AbCphZ* and *PaAotO* were subcloned into vector pBacP, a pJ411-derived plasmid that encodes a C-terminal, tobacco etch virus (TEV) protease cleavable, 8x histidine affinity tag. All cloning and mutagenesis steps were accomplished *in cellulo*, transforming PCR fragments containing overlapping ends into DH5- $\alpha$  *E. coli* cells. Protein expression was carried out in *E. coli* BL21(DE3) grown in TB media supplemented with 150  $\mu$ g/mL kanamycin. Cultures were grown at 37 °C until an OD<sub>600</sub> of ~1 was reached, at which time the growth temperature was lowered to 18 °C. Protein expression was induced by the addition of 0.2 mM isopropyl  $\beta$ -d-1-thiogalactopyranoside (IPTG), and culture grown for ~20 hours prior to harvesting by centrifugation. All following steps were carried out at 4 °C. The cell pellets were resuspended in buffer A (250 mM NaCl, 50 mM Tris pH 8.0, 10 mM imidazole, 2 mM  $\beta$ -mercaptoethanol) supplemented with crystals of lysozyme and DNase I and lysed on ice by sonication. The lysate was clarified by centrifugation at 40,000 g for 30 minutes and then applied onto a HisTrap HP column (Cytiva, USA). Following loading, the protein was washed with at least 20 column volumes of buffer B (buffer A with 30 mM imidazole) and eluted with buffer C (buffer A with 250 mM imidazole). For structural studies, the protein was dialyzed overnight against buffer D (250 mM NaCl, 20 mM Tris pH 8.0, 5 mM  $\beta$ -mercaptoethanol) in the presence of TEV protease for removal of the octahistidine tag. Following tag cleavage, the protein was again applied to a HisTrap column and the flow through was collected. For all protein preparations, sample was concentrated using an Amicon<sup>®</sup> Ultra Centrifugal filter and applied to a Superdex200 16/60 column (Cytiva, USA) equilibrated in buffer E (100 mM NaCl, 20 mM Tris pH 8.0, 1 mM dithiothreitol). Fractions containing the highest protein purity were pooled, and following concentration were supplemented with glycerol to a final volume of 15% and flash frozen in liquid nitrogen for storage.

### Protein crystallization, data collection, structure solution and refinement

*AbCphZ* crystals were grown using the sitting drop method using a reservoir volume of 500  $\mu$ l and a drop containing 2  $\mu$ l of protein sample in buffer E and 2  $\mu$ l of crystallization buffer. The crystallization buffer for WT *AbCphZ* (5 mg/mL) contained 0.1 M bis-tris propane pH 7.5, 24% PEG3350, 0.2 M NaBr, 10 mM spermine and 10 mM LiCl. Crystals were grown at 4 °C, dehydrated by allowing the drop to equilibrate through vapor diffusion against 0.1 M bis-tris propane pH 7.5, 30% PEG3350, 0.2 M NaBr and 10 mM spermine overnight, and then cryo-protected by dipping them in 0.1 M bis-tris propane pH 7.5, 20% PEG3350, 20% ethylene glycol, 0.2 M NaBr and 10 mM spermine. Data were collected at the Advanced Photon Source (APS) beamline 24-ID-C. The structure was solved by molecular replacement using PDB codes 3FMC and 3LWU as search models. The crystals of E251A *AbCphZ* were grown as above and soaked in a cryo-protection solution containing 10 mM  $\beta$ -Asp-Arg for 30 minutes prior to freezing. Data were collected at the Advanced Light Source (ALS) beamline 5.0.1. All datasets were both indexed and integrated in DIALS(66) and merged in Aimless(67) implemented in CCP4i2 suite(68), and in parallel indexed, integrated and merged in HKL2000(69). The structures were refined in REFMAC(70), Rosetta(71), PHENIX (72) and Coot(73). Figures were prepared in PyMOL (Schrödinger, Inc.). For anomalous difference map calculation, data from E251A *AbCphZ* crystals were collected at the Canadian Light Source (CLS) beamline CMCF-BM at a wavelength of 1.283 Å, processed in DIALS (66), Aimless (67), retaining the anomalous information. The anomalous difference map was calculated using these data in the program CCP4i2(68).

#### Enzyme activity assays

An Asp release assay(36) was used to detect Asp release from  $\beta$ -Asp-X dipeptides: Each 100  $\mu$ l reaction contained 100 mM HEPES pH 8.2, 20 mM KCl, 5 mM  $\alpha$ -ketoglutarate, 500 nM purified enzyme, 0.3 U malate dehydrogenase, 1 mM NADH, 2.4 U aspartate aminotransferase and 1 mM dipeptide substrate. For other Arg containing substrates, Arg release was monitored by using a free Arg detection kit (K-LARGE, Neogen, USA): The 135  $\mu$ l reactions contained 15  $\mu$ l buffer solution, 10  $\mu$ l NADPH solution, 1  $\mu$ l GIDH suspension, 2.5  $\mu$ l urease solution, 1  $\mu$ l arginase suspension, 1 mM substrate and 500 nM enzyme. For both assays, reaction progression was monitored by following 340 nm transmittance in 96-well plates. Data were collected using a SpectraMax Paradigm (Molecular Devices, USA) and analyzed using GraphPad Prism (GraphPad, USA).  $\beta$ -Asp-Arg dipeptides were purified as previously described(26).  $\beta$ -Asp-Ala and  $\alpha$ -Asp-Arg

were purchased from Bachem (Switzerland), N<sup>2</sup>-acetyl arginine, β-Asp-Lys and β-Asp-Leu from Toronto Research Chemicals (Canada), β-Asp-Asp from AchemBlock (USA), and N<sup>2</sup>-succinyl arginine from BLD Pharmatech (USA).

#### Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a MicroCal iTC200 (GE Healthcare, USA) at 25 °C. For β-Asp-Arg, the cell contained 200 μM purified E251A AbCphZ in buffer F (100 mM NaCl, 20 mM Tris pH 8.0, 5 mM β-mercaptoethanol) and the syringe contained 1 mM β-Asp-Arg in buffer F. For L-Arg and L-Asn, the cell contained 400 μM purified E251A AbCphZ and the syringe contained 10 mM substrate, both in buffer F. A total of nineteen 2 μl injections were interspaced by 180 seconds each. Data were analyzed using Microcal Origin 7.0 (OriginLab, USA) with a binding model stoichiometry of 1:1. Measurements were performed in triplicates.

#### Inductively coupled plasma mass spectrometry (ICP-MS)

AbCphZ was buffer-exchanged into 100 mM (NH<sub>4</sub>)<sub>2</sub>CO<sub>3</sub> by performing gel filtration chromatography with a Superdex S200 10/300 column equilibrated with that buffer. The protein was concentrated to 100 μM and analyzed by ICP-MS at the Center for Applied Isotope Studies, University of Georgia. A sample of the buffer eluted from the column was used as a control.

#### Bioinformatic analyses

For bioinformatic analyses, protein sequences were considered homologous if above a threshold of >70% coverage and >30% identity. This threshold almost certainly discards bone fide homologs, but is important to exclude false positives: A set cyanobacterial sequences featuring 21-28% identity and 83-87% coverage to AbCphZ is returned in BLAST searches if lower thresholds are used (e.g. E value < 0.05). These cyanobacterial sequences contain residues for binding the Zn<sup>2+</sup> and the Arg moiety but not the Asp portion of β-Asp-Arg, and thus cannot be assumed to be isofunctional with CphZ. The >70% coverage and >30% identity threshold appears effective in limiting false positives, as measured by the conservation of active site residues: i) Zn<sup>2+</sup> binding residues: H50 99.7% ; E53 99.7%; H179 100%; ii) Nucleophilic water activation: E251 99.96%;

iii) Asp moiety binding: D181 99.4%; K366 97.7% as K, H or R; iv) Arg moiety binding: R102 99.96%; N112: 99.98%; R113 100%; S216 93.2%; D222 99.8%; E223 99.96%.

For bioinformatic analyses of gene co-occurrence, all complete bacterial genomes in the NCBI RefSeq database were used to create a local database (May 2022). Cblaster(58) was used in local mode to search this database using several templates for CphA1 (accession codes WP\_028947105.1, WP\_004925893.1 and WP\_015942562.1), cyanophycinase (WP\_011058003.1, WP\_004925892.1 and Q8KQN8.1) and *PaAotO* and *AbCphZ*. The resulting binary table was analyzed using Excel.

### Bacterial strains

The *P. aeruginosa* strains used in this study are listed in SI Appendix, Table S4. The *aotO*, *aotP* and *aotM* transposon mutants were obtained from the PAO1 two-allele transposon mutant library and confirmed by PCR(50, 51) and sequencing.

### *P. aeruginosa* growth conditions

*P. aeruginosa* was cultured in minimal media (22 mM KH<sub>2</sub>PO<sub>4</sub>, 42 mM Na<sub>2</sub>HPO<sub>4</sub>, 8.6 mM NaCl, 2.0 mM MgSO<sub>4</sub>, 5.0 μM FeSO<sub>4</sub>, pH 7.0) supplemented with carbon or nitrogen source as indicated in each experiment, as per Lundgren et al. (74). Subcultures were inoculated from agar plates (LB agar for WT, LB agar with 50 μg/ml tetracycline for Tc-R mutants) and grown in 5 ml minimal media supplemented with 20 mM D-glucose and 10 mM NH<sub>4</sub>Cl, in 16 mm x 150 mm sterile glass tubes for 6 h (230 rpm, 37 °C). Following subculture, the bacteria were diluted to an OD<sub>600</sub> of 0.05 in 5 ml of fresh minimal media supplemented with 20 mM D-glucose and 10 mM NH<sub>4</sub>Cl and grown for 24 h (230 rpm, 37 °C). Cells were harvested by centrifugation (7800xg for 5 min), washed twice in 5 ml of minimal media with no supplementation to remove all residual nutrients, and then resuspended in minimal media at an OD<sub>600</sub> of ~0.05. The final culture media were then supplemented with one of the following: 20 mM D-glucose and 10 mM NH<sub>4</sub>Cl, 10 mM L-aspartate, 10 mM β-Asp-Arg, 10 mM β-Asp-Ala or 10 mM L-arginine as indicated. For growth experiments, bacteria were incubated in either 15 ml (in 150 ml flasks, 230 rpm, 37°C) or 200 μl (in 96-well plates, continuous shaking 282 CPM (3 mm), 37°C) using a BioTek Cytation 5 as indicated. Bacterial growth was monitored by OD<sub>600</sub> (using a GENESYS UV scanning

spectrophotometer or BioTek Cytation 5 plate reader) and by viable bacterial counts (using standard serial dilution and plating on LB agar).

### **Data availability**

The protein structures solved in this study have been deposited to the Protein Data Bank, under the accession codes 8EIN (*AbCphZ*) and 8EIP (*AbCphZ\_E251A* +  $\beta$ -Asp-Arg).

### **Acknowledgements**

We thank all the members of the Schmeing and Nguyen labs for important advice and ongoing discussions on this project, Kim Munro for help with ITC experiments, Nancy Rogerson for proofreading, and synchrotron staff D. Neau (Advanced Photon Source), M. Allaire (Advanced Light Source) and CLS staff for facilitating remote collection of diffraction datasets. This study includes work based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Beamline 5.0.1 of the Advanced Light Source, a U.S. DOE Office of Science User Facility under Contract No. DE-AC02-05CH11231, is supported in part by the ALS-ENABLE program funded by the National Institutes of Health, National Institute of General Medical Sciences, grant P30 GM124169-01. Part of the research described in this paper was performed using beamline CMCF-BM at the Canadian Light Source, a national research facility of the University of Saskatchewan, which is supported by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council (NSERC), the National Research Council (NRC), the Canadian Institutes of Health Research (CIHR), the Government of Saskatchewan, and the University of Saskatchewan. This work was funded by CIHR Project Grant 178084 and a Canada Research Chair to TMS, and by NSERC Discovery Grant RGPIN-2019-06408 to DN.

### **Competing interests**

The authors declare no competing interests.

## Figures

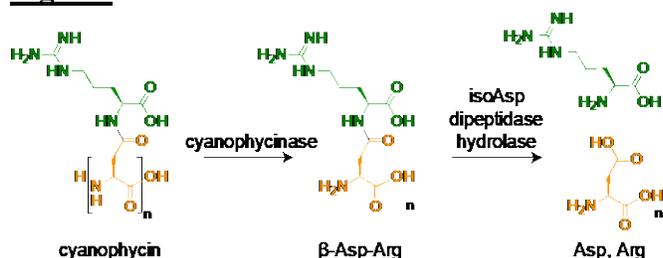
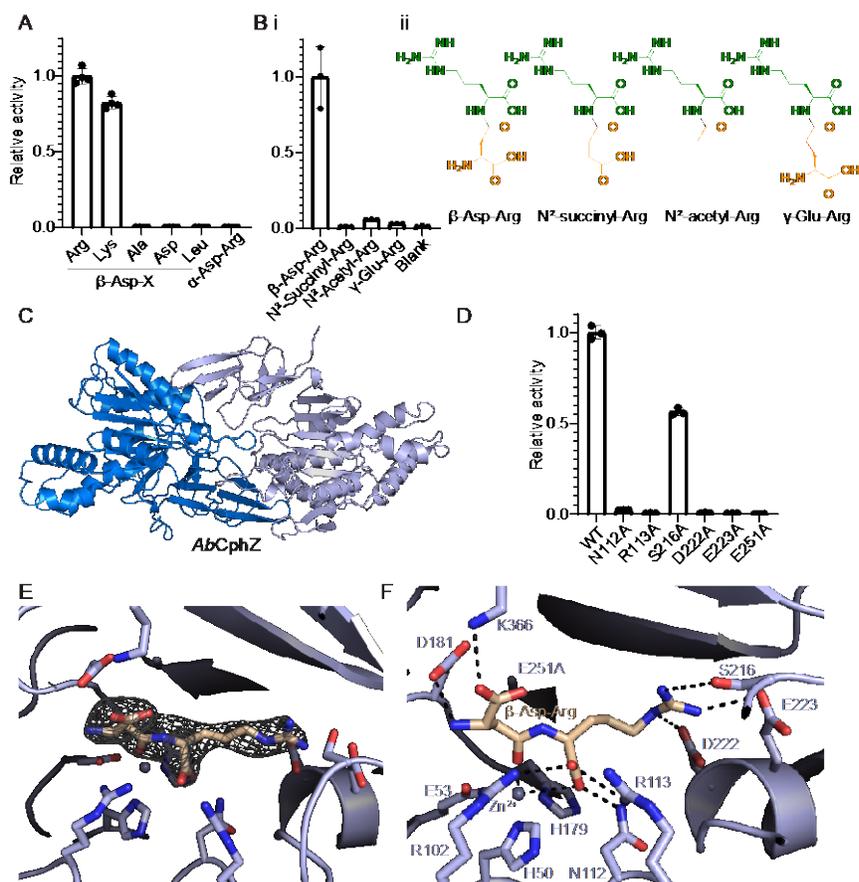
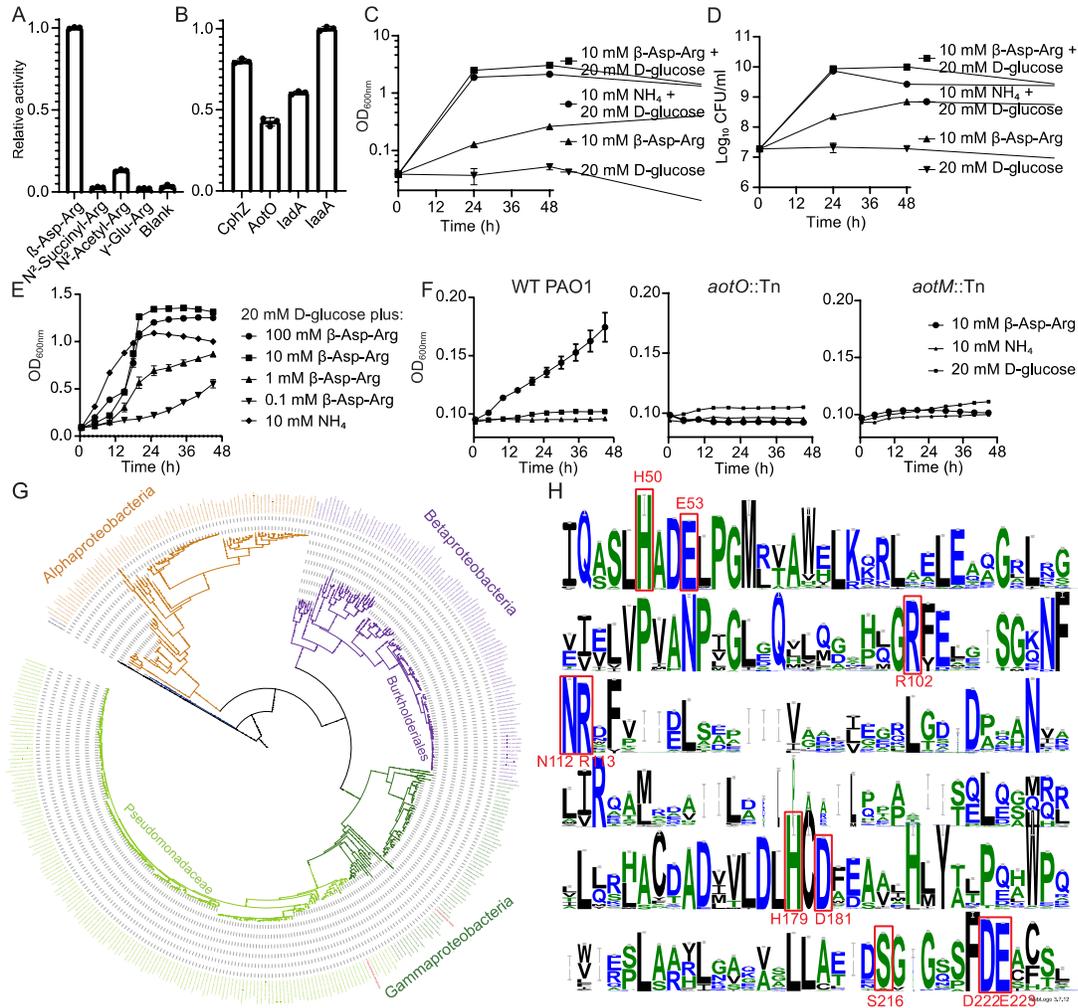


Figure 1. The two steps of cyanophycin biodegradation.



**Figure 2. Structural and biochemical characterization of *AbCphZ*.** A. Asp release assay of *AbCphZ* with various Asp-dipeptides. *AbCphZ* displays high specificity towards  $\beta$ -Asp-Arg/Lys dipeptides. Error bars represent the standard deviation of the mean of  $n=4$  replicates. (B) Arg release assay of *AbCphZ* with Arg-containing substrates. *AbCphZ* preferentially cleaves  $\beta$ -Asp-Arg. Error bars represent the standard deviation of the mean of  $n=3$  replicates. (C) Overview of the crystal structure of *AbCphZ*, which is a dimer. (D) Asp release assay of *AbCphZ* mutants with  $\beta$ -Asp-Arg. The highly conserved N112, R113, D222, E223 and E251 are essential for activity, while mutation of S216, which is not as highly conserved (Fig 3F), decreases activity to approximately half. Error bars represent the standard deviation of the mean of  $n=3$  replicates. (E) Polder electron density map showing unbiased density for the substrate  $\beta$ -Asp-Arg bound to *AbCphZ* E251A. The map is displayed at contour at  $4.5 \sigma$  and carved to within 4 Å of  $\beta$ -Asp-Arg. (F) Detailed view of the active site showing the substrate  $\beta$ -Asp-Arg bound to E251A, coordinated by a zinc ion (Zn<sup>2+</sup>) and other residues (D181, K366, S216, E223, D222, R113, H179, N112, H150, R102, E3).

(F) Detailed view of the active site of *AbCphZ* E251A with bound  $\beta$ -Asp-Arg. Note that the E251A mutant retains zinc, but is unable to position an activated water for nucleophilic attack, as also observed in other zinc proteases (44).



**Figure 3. Cyanophycin dipeptide hydrolase phylogeny and role in *P. aeruginosa*.** (A) Arg release assay with Arg-containing substrates. *PaAotO* preferentially cleaves  $\beta$ -Asp-Arg. Error bars represent the standard deviation of the mean of n=3 replicates. (B) Arg release assay of *AbCphZ*, *PaAotO* and two isoaspartyl dipeptidase enzymes described in a previous study(40) with the substrate  $\beta$ -Asp-Arg. Error bars represent the standard deviation of the mean of n=3 replicates. (C) Growth of *P. aeruginosa* wild-type PAO1 in minimal media supplemented with 20 mM D-glucose and 10 mM  $\text{NH}_4\text{Cl}$  (●), 20 mM D-glucose and 10 mM  $\beta$ -Asp-Arg (■), 10 mM  $\beta$ -Asp-Arg (▲) or 20 mM D-glucose (▼). (D) Viable cell count (log CFU/ml) of PAO1 WT grown in minimal media supplemented with 20 mM D-glucose and 10 mM  $\text{NH}_4\text{Cl}$  (●), 20 mM D-glucose and 10 mM  $\beta$ -Asp-Arg (■), 10 mM  $\beta$ -Asp-Arg (▲) or 20 mM D-glucose (▼). Cell counts in media containing  $\beta$ -Asp-Arg increased from  $1.9 \times 10^7$  (SD=0.37) CFU/ml at t = 0 h to  $5.7 \times 10^8$  (SD=0.21) CFU/ml at t = 48 h. Data points are from a representative experiment; mean values (n=3)  $\pm$  SD. The experiment was repeated twice. (E) Growth of *P. aeruginosa* PAO1 in 200  $\mu\text{l}$  minimal media in 96 well microplate supplemented with 20 mM D-glucose and 100 mM  $\beta$ -Asp-Arg (●), 20 mM D-glucose and 10 mM  $\beta$ -Asp-Arg (■), 20 mM D-glucose and 1 mM  $\beta$ -Asp-Arg (▲) or 20 mM D-

glucose and 0.1 mM  $\beta$ -Asp-Arg ( $\blacktriangledown$ ). Growth with 20 mM D-glucose and 10 mM  $\text{NH}_4\text{Cl}$  ( $\blacklozenge$ ) is also shown. Growth is saturated at 100 mM  $\beta$ -Asp-Arg, likely due to limitation of available glucose. Data points represent mean values ( $n = 6$ )  $\pm$  SEM from two independent experiments. (F) Growth of *P. aeruginosa* PAO1 WT, *aotM*::Tn, and *aotO*::Tn mutants in 200  $\mu\text{L}$  minimal media in 96 well microplate supplemented with 10 mM  $\text{NH}_4\text{Cl}$  ( $\blacktriangle$ ), 20 mM D-glucose ( $\blacksquare$ ), or 10 mM  $\beta$ -Asp-Arg ( $\bullet$ ). Data points represent mean values ( $n = 6$ )  $\pm$  SEM from two independent experiments. (G) Phylogenetic tree of CphZ/AotO distribution of CphZ proteins encoded by for completely sequenced bacterial genomes (53). *AbCphZ* and *PaAotO* are colored in red. The blue and black colored entries are deltaproteobacterial and mycobacterial respectively, presumably resulting from horizontal gene transfer. (H) A Weblogo (56) of the *AbCphZ* region 45-227 shows high conservation of active site and substrate binding residues in the cyanophycin dipeptide hydrolases. E251 (not shown) is also very highly conserved, while K366 can be K, H or R.

## Tables

**Table 1. Co-occurrence and clustering of *cphA1*, cyanophycinase genes and *cphZ***

	Total	<i>cphA1</i>	cyano- phycinase	<i>cphZ</i>	<i>cphA1</i> + <i>cphZ</i> + cyanophycinase	<i>cphA1</i> + cyanophycinase	<i>cphZ</i> + cyanophycinase	<i>cphA1</i> + <i>cphZ</i>
In genome	3095	1614	840	1364	41	659	50	56
Clustered	563	543	562	33	12	542	32	13

## References

1. I. Sharon *et al.*, Structures and function of the amino acid polymerase cyanophycin synthetase. *Nat Chem Biol* **17**, 1101–1110 (2021).
2. G. Fuser, A. Steinbuchel, Analysis of genome sequences for genes of cyanophycin metabolism: identifying putative cyanophycin metabolizing prokaryotes. *Macromol Biosci* **7**, 278–296 (2007).
3. M. Burnat, A. Herrero, E. Flores, Compartmentalized cyanophycin metabolism in the diazotrophic filaments of a heterocyst-forming cyanobacterium. *Proc Natl Acad Sci U S A* **111**, 3823–3828 (2014).
4. S. Picossi, A. Valladares, E. Flores, A. Herrero, Nitrogen-regulated genes for the metabolism of cyanophycin, a bacterial nitrogen reserve polymer: expression and mutational analysis of two cyanophycin synthetase and cyanophycinase gene clusters in heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. *J Biol Chem* **279**, 11582–11592 (2004).
5. R. D. Simon, The biosynthesis of multi-L-arginyl-poly(L-aspartic acid) in the filamentous cyanobacterium *Anabaena cylindrica*. *Biochim Biophys Acta* **422**, 407–418 (1976).
6. M. M. Allen, P. J. Weathers, Structure and composition of cyanophycin granules in the cyanobacterium *Aphanocapsa* 6308. *J Bacteriol* **141**, 959–962 (1980).
7. A. H. Mackerras, N. M. de Chazal, G. D. Smith, Transient accumulations of cyanophycin in *Anabaena cylindrica* and *Synechocystis* 6308. *Microbiology* **136**, 2057–2065 (1990).
8. K. Zou *et al.*, Cyanophycin granule polypeptide: a neglected high value-added biopolymer, synthesized in activated sludge on a large scale. *Appl Environ Microbiol* 10.1128/aem.00742-22, e0074222 (2022).
9. A. S. Sindi *et al.*, Effect of a reduced fat and sugar maternal dietary intervention during lactation on the infant gut microbiome. *Front Microbiol* **13**, 900702 (2022).
10. A. Borzi, Le comunicazioni intracellulari delle nostochinee. *Malpighia - Rassegna mensile di botanica* **1** (1886-1887).
11. R. D. Simon, Cyanophycin granules from the blue-green alga *anabaena cylindrica*: a reserve material consisting of copolymers of aspartic acid and arginine. *Proc Natl Acad Sci U S A* **68**, 265–267 (1971).
12. S. Liotenberg, D. Campbell, R. Rippka, J. Houmard, N. T. de Marsac, Effect of the nitrogen source on phycobiliprotein synthesis and cell reserves in a chromatically adapting filamentous cyanobacterium. *Microbiology* **142**, 611–622 (1996).
13. B. Liang *et al.*, Cyanophycin mediates the accumulation and storage of fixed carbon in non-heterocystous filamentous cyanobacteria from coniform mats. *PLoS One* **9**, e88142 (2014).
14. L. L. Wingard *et al.*, Cyanophycin production in a phycoerythrin-containing marine *synechococcus* strain of unusual phylogenetic affinity. *Appl Environ Microbiol* **68**, 1772–1777 (2002).
15. J. A. Finzi-Hart *et al.*, Fixation and fate of C and N in the cyanobacterium *Trichodesmium* using nanometer-scale secondary ion mass spectrometry. *Proc Natl Acad Sci U S A* **106**, 6345–6350 (2009).

16. A. Sukenik *et al.*, Carbon assimilation and accumulation of cyanophycin during the development of dormant cells (akinetes) in the cyanobacterium *Aphanizomenon ovalisporum*. *Front Microbiol* **6** (2015).
17. M. Maheswaran, K. Ziegler, W. Lockau, M. Hagemann, K. Forchhammer, PII-regulated arginine synthesis controls accumulation of cyanophycin in *Synechocystis* sp. strain PCC 6803. *J Bacteriol* **188**, 2730-2734 (2006).
18. K. Ziegler, D. P. Stephan, E. K. Pistorius, H. G. Ruppel, W. Lockau, A mutant of the cyanobacterium *Anabaena variabilis* ATCC 29413 lacking cyanophycin synthetase: growth properties and ultrastructural aspects. *FEMS Microbiol Lett* **196**, 13–18 (2001).
19. H. Li, D. M. Sherman, S. Bao, L. A. Sherman, Pattern of cyanophycin accumulation in nitrogen-fixing and non-nitrogen-fixing cyanobacteria. *Arch Microbiol* **176**, 9–18 (2001).
20. A. Omairi-Nasser, V. Mariscal, J. R. Austin, R. Haselkorn, Requirement of Fra proteins for communication channels between cells in the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp PCC 7120. *Proc Natl Acad Sci U S A* **112**, E4458–E4464 (2015).
21. K. Ziegler *et al.*, Molecular characterization of cyanophycin synthetase, the enzyme catalyzing the biosynthesis of the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartate (cyanophycin). *Eur J Biochem* **254**, 154–159 (1998).
22. I. Sharon *et al.*, A cryptic third active site in cyanophycin synthetase creates primers for polymerization. *Nat Commun* **13**, 3923 (2022).
23. M. Frommeyer, A. Steinbuchel, Increased lysine content is the main characteristic of the soluble form of the polyamide cyanophycin synthesized by recombinant *Escherichia coli*. *Appl Environ Microbiol* **79**, 4474–4483 (2013).
24. M. Frommeyer, L. Wiefel, A. Steinbuchel, Features of the biotechnologically relevant polyamide family "cyanophycins" and their biosynthesis in prokaryotes and eukaryotes. *Crit Rev Biotechnol* **36**, 153–164 (2016).
25. F. Klemke *et al.*, CphA2 is a novel type of cyanophycin synthetase in N<sub>2</sub>-fixing cyanobacteria. *Microbiology* **162**, 526–536 (2016).
26. I. Sharon, M. Grogg, D. Hilvert, T. M. Schmeing, Structure and function of the  $\beta$ -Asp-Arg polymerase cyanophycin synthetase 2. *ACS Chem Biol* **17** (2022).
27. R. Richter, M. Hejazi, R. Kraft, K. Ziegler, W. Lockau, Cyanophycinase, a peptidase degrading the cyanobacterial reserve material multi-L-arginyl-poly-L-aspartic acid (cyanophycin): molecular cloning of the gene of *Synechocystis* sp. PCC 6803, expression in *Escherichia coli*, and biochemical characterization of the purified enzyme. *Eur J Biochem* **263**, 163–169 (1999).
28. I. Sharon, M. Grogg, D. Hilvert, T. M. Schmeing, The structure of cyanophycinase in complex with a cyanophycin degradation intermediate. *Biochim Biophys Acta Gen Subj* **1866**, 130217 (2022).
29. M. Hejazi *et al.*, Isoaspartyl dipeptidase activity of plant-type asparaginases. *Biochem J* **364**, 129–136 (2002).
30. J. D. Lowenson, E. Kim, S. G. Young, S. Clarke, Limited accumulation of damaged proteins in l-isoaspartyl (D-aspartyl) O-methyltransferase-deficient mice. *J Biol Chem* **276**, 20695–20702 (2001).
31. J. L. Radkiewicz, H. Zipse, S. Clarke, K. N. Houk, Accelerated racemization of aspartic acid and asparagine residues via succinimide intermediates: an *ab initio* theoretical exploration of mechanism. *J Am Chem Soc* **118**, 9148-9155 (1996).

32. D. W. Aswad, M. V. Paranandi, B. T. Schurter, Isoaspartate in peptides and proteins: formation, significance, and analysis. *J Pharm Biomed Anal* **21**, 1129–1136 (2000).
33. J. D. Gary, S. Clarke, Purification and characterization of an isoaspartyl dipeptidase from *Escherichia coli*. *J Biol Chem* **270**, 4076–4087 (1995).
34. E. Kim, J. D. Lowenson, D. C. MacLaren, S. Clarke, S. G. Young, Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice. *Proc Natl Acad Sci U S A* **94**, 6132–6137 (1997).
35. E. E. Haley, Purification and properties of a  $\beta$ -aspartyl peptidase from *Escherichia coli*. *J Biol Chem* **243**, 5748–5752 (1968).
36. R. Marti-Arbona *et al.*, Mechanism of the reaction catalyzed by isoaspartyl dipeptidase from *Escherichia coli*. *Biochemistry* **44**, 7115–7124 (2005).
37. D. Borek *et al.*, Expression, purification and catalytic activity of *Lupinus luteus* asparagine beta-amidohydrolase and its *Escherichia coli* homolog. *Eur J Biochem* **271**, 3215–3226 (2004).
38. A. Prahl, M. Pazgier, M. Hejazi, W. Lockau, J. Lubkowski, Structure of the isoaspartyl peptidase with L-asparaginase activity from *Escherichia coli*. *Acta Crystallogr D Biol Crystallogr* **60**, 1173–1176 (2004).
39. E. Flores, S. Arévalo, M. Burnat, Cyanophycin and arginine metabolism in cyanobacteria. *Algal Research* **42**, 101577 (2019).
40. I. Sharon, T. M. Schmeing, Characterization of promiscuous isoaspartyl dipeptidases that catalyze the final step in cyanophycin degradation. *Manuscript in preparation* (2022).
41. B. L. Schneider, A. K. Kiupakis, L. J. Reitzer, Arginine catabolism and the arginine succinyltransferase pathway in *Escherichia coli*. *J Bacteriol* **180**, 4278–4286 (1998).
42. D. C. Rees, M. Lewis, W. N. Lipscomb, Refined crystal structure of carboxypeptidase A at 1.54 Å resolution. *J Mol Biol* **168**, 367–387 (1983).
43. J. H. Cho *et al.*, Insight into the stereochemistry in the inhibition of carboxypeptidase A with N-(hydroxyaminocarbonyl)phenylalanine: binding modes of an enantiomeric pair of the inhibitor to carboxypeptidase A. *Bioorg Med Chem* **10**, 2015–2022 (2002).
44. N. Cerda-Costa, F. X. Gomis-Ruth, Architecture and function of metallopeptidase catalytic domains. *Protein Sci* **23**, 123–144 (2014).
45. I. L. Alberts, K. Nadassy, S. J. Wodak, Analysis of zinc binding sites in protein crystal structures. *Protein Sci* **7**, 1700–1716 (1998).
46. A. Marchler-Bauer *et al.*, CDD: NCBI's conserved domain database. *Nucleic Acids Res* **43**, D222–226 (2015).
47. T. Nishijyo, S. M. Park, C. D. Lu, Y. Itoh, A. T. Abdelal, Molecular characterization and regulation of an operon encoding a system for transport of arginine and ornithine and the ArgR regulatory protein in *Pseudomonas aeruginosa*. *J Bacteriol* **180**, 5559–5566 (1998).
48. C. D. Lu, Z. Yang, W. Li, Transcriptome analysis of the ArgR regulon in *Pseudomonas aeruginosa*. *J Bacteriol* **186**, 3855–3861 (2004).
49. D. A. Johnson *et al.*, High-throughput phenotypic characterization of *Pseudomonas aeruginosa* membrane transport genes. *PLoS Genet* **4**, e1000211 (2008).
50. D. Nguyen *et al.*, Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. *Science* **334**, 982–986 (2011).
51. K. Held, E. Ramage, M. Jacobs, L. Gallagher, C. Manoil, Sequence-verified two-allele transposon mutant library for *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **194**, 6387–6389 (2012).

52. L. Zimmermann *et al.*, A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol* **430**, 2237-2243 (2018).
53. N. A. O'Leary *et al.*, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research* **44**, D733-D745 (2016).
54. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**, W256–W259 (2019).
55. D. H. Parks *et al.*, GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Research* **50**, D785-D794 (2021).
56. G. E. Crooks, G. Hon, J. M. Chandonia, S. E. Brenner, WebLogo: a sequence logo generator. *Genome Res* **14**, 1188–1190 (2004).
57. M. Johnson *et al.*, NCBI BLAST: a better web interface. *Nucleic Acids Res* **36**, W5–9 (2008).
58. C. L. M. Gilchrist *et al.*, cblaster: a remote search tool for rapid identification and visualization of homologous gene clusters. *Bioinformatics Advances* **1**, vbab016 (2021).
59. M. Nikolaidis, D. Mossialos, S. G. Oliver, G. D. Amoutzias, Comparative analysis of the core proteomes among the *Pseudomonas* major evolutionary groups reveals species-specific adaptations for *Pseudomonas aeruginosa* and *Pseudomonas chlororaphis*. *Diversity* **12** (2020).
60. M. Obst, F. B. Oppermann-Sanio, H. Luftmann, A. Steinbuchel, Isolation of cyanophycin-degrading bacteria, cloning and characterization of an extracellular cyanophycinase gene (*cphE*) from *Pseudomonas anguilliseptica* strain BI. The *cphE* gene from *P. anguilliseptica* BI encodes a cyanophycin hydrolyzing enzyme. *J Biol Chem* **277**, 25096–25105 (2002).
61. M. Obst, A. Sallam, H. Luftmann, A. Steinbuchel, Isolation and characterization of gram-positive cyanophycin-degrading bacteria-kinetic studies on cyanophycin depolymerase activity in aerobic bacteria. *Biomacromolecules* **5**, 153–161 (2004).
62. M. Obst, A. Krug, H. Luftmann, A. Steinbuchel, Degradation of cyanophycin by *Sedimentibacter hongkongensis* strain KI and *Citrobacter amalonaticus* strain G isolated from an anaerobic bacterial consortium. *Appl Environ Microbiol* **71**, 3642–3652 (2005).
63. A. Sallam, A. Steinbuchel, Anaerobic and aerobic degradation of cyanophycin by the denitrifying bacterium *Pseudomonas alcaligenes* strain DIP1 and role of three other coisolates in a mixed bacterial consortium. *Appl Environ Microbiol* **74**, 3434–3443 (2008).
64. A. Sallam, A. Steinbuchel, Cyanophycin-degrading bacteria in digestive tracts of mammals, birds and fish and consequences for possible applications of cyanophycin and its dipeptides in nutrition and therapy. *J Appl Microbiol* **107**, 474–484 (2009).
65. J. P. Bourdineaud *et al.*, Characterization of the arcD arginine:ornithine exchanger of *Pseudomonas aeruginosa*. Localization in the cytoplasmic membrane and a topological model. *J Biol Chem* **268**, 5417-5424 (1993).
66. J. Beilsten-Edmands *et al.*, Scaling diffraction data in the DIALS software package: algorithms and new approaches for multi-crystal scaling. *Acta Crystallogr D Struct Biol* **76**, 385–399 (2020).
67. P. R. Evans, G. N. Murshudov, How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* **69**, 1204–1214 (2013).

68. L. Potterton *et al.*, CCP4i2: the new graphical user interface to the CCP4 program suite. *Acta Crystallogr D Struct Biol* **74**, 68–84 (2018).
69. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326 (1997).
70. O. Kovalevskiy, R. A. Nicholls, G. N. Murshudov, Automated refinement of macromolecular structures at low resolution using prior information. *Acta Crystallogr D Struct Biol* **72**, 1149–1161 (2016).
71. Y. Song *et al.*, High-resolution comparative modeling with RosettaCM. *Structure* **21**, 1735–1742 (2013).
72. P. D. Adams *et al.*, The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).
73. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132 (2004).
74. B. R. Lundgren *et al.*, Utilization of L-glutamate as a preferred or sole nutrient in *Pseudomonas aeruginosa* PAO1 depends on genes encoding for the enhancer-binding protein AauR, the sigma factor RpoN and the transporter complex AatJQMP. *Bmc Microbiol* **21** (2021).