
DOI link
https://doi.org/10.1074/jbc.TM117.000180

ePrints link
http://eprint.ncl.ac.uk/247333

Date deposited
13/04/2018

Copyright
This research was originally published in the Journal of Biological Chemistry. © 2018 The American Society for Biochemistry and Molecular Biology
ABSTRACT
Copper is essential for most organisms as a cofactor for key enzymes involved in fundamental processes such as respiration and photosynthesis. However, copper also has toxic effects in cells, which is why eukaryotes and prokaryotes have evolved mechanisms for safe copper handling. A new family of bacterial proteins use a Cys-rich four-helix bundle to safely store large quantities of Cu(I). The work leading to the discovery of these proteins, their properties and physiological functions, and how their presence potentially impacts on current views of bacterial copper handling and use are discussed in this review.

Copper in biology
The utilization of metals by biological systems is highly paradoxical. On the one hand, metal ions provide proteins access to chemistry that would otherwise be impossible using the organic reactions that can be catalysed by amino acid side chains. On the other hand, many of these metal ions can be toxic to cells. Copper is essential for most organisms as the cofactor for key enzymes involved in important processes such as respiration and photosynthesis (1-7). Ideas about the cellular toxicity of copper have developed in recent years, from solely being attributed to the generation of reactive oxygen species (ROS) (8-11). An emerging mechanism appears to be driven by the ability of copper to bind tightly at the active sites of metalloenzymes, and particularly those containing iron-sulfur clusters. This not only destroys the reactivity of the mis-metallated protein, but releases Fe that can produce ROS (9). This toxicity is the reason why aquated (“free”) copper ions should not exist in cells, and that copper is predicted to be highly restricted in eukaryotes (12) and prokaryotes (13). Copper availability appears to be largely constrained by the use of high-affinity sites in proteins (12-14), although “pools” of copper bound by other molecules are important (4, 5, 11, 15-18).

Approaches used by cells to enable safe copper handling, referred to as copper homeostasis, include sensors, transporters, chaperones and insertion proteins with high affinity and specificity for copper (3-7, 12-14, 19-22). A well-characterized family of copper-homeostasis proteins are the copper-transporting P-type ATPases, which can remove this metal ion from the cytosol (4-7, 20-24). In eukaryotes, these copper-efflux pumps work with a cytosolic copper metallochaperone (ATOX1 in humans and Atx1 in yeast) to facilitate import into the trans-Golgi network for secreted copper enzymes (4, 5, 19, 24, 25). The two Cu-ATPases in humans (ATP7A and ATP7B) can relocate to the plasma membrane to remove excess intracellular copper when necessary (4, 24). In bacteria, the production of the copper-efflux pump CopA (23) is controlled by transcriptional regulators (sensors) such as CueR (13) and CsoR (26). CopA can work either alone or in concert with the ATOX1/Atx1 homologue CopZ to remove cytosolic copper (5-7, 20-23, 27, 28). It has recently been found that in bacteria not previously thought to possess this copper metallochaperone such as Escherichia coli, CopZ can be made from the CopA gene by “programmed ribosomal frameshifting” (29).
It is emerging that the human immune system uses the toxicity of copper to attack invading pathogens. Previous minireviews in the “Thematic Series on Metals in Biology” have discussed copper biochemistry (3, 22), emphasising its role in pathogenicity (30-34). We will therefore only touch on this issue briefly towards the end of our minireview. The main topic here is the recently discovered ability of bacteria to safely store copper using a highly novel approach (35). The more widespread and abundant class of the new family of bacterial proteins that can perform this function are cytosolic (36). This is somewhat controversial, as a widely accepted view is that bacteria have evolved not to use cytosolic copper enzymes as a way to help avoid the potential toxicity associated with their metalation (6, 13, 37).

Discovery of a new bacterial copper storage protein and its characterization

Eukaryotes are able to store cytosolic copper using metallothioneins (MTs) (38-41). Related proteins have been characterized in pathogenic mycobacteria (42), but the idea that bacterial copper storage systems could be more common was unknown. This changed with the discovery of a new family of copper storage proteins, the Csps, in the methane-oxidizing bacterium (methanotroph) Methylosinus trichosporium OB3b (35). It is not surprising that such a finding about copper biochemistry was made in methanotrophs as these Gram-negative organisms use large amounts of copper to metabolise methane via the membrane-bound (particulate) methane monooxygenase (pMMO). This enzyme catalyzes the conversion of methane to methanol in almost all methanotrophs (17). pMMO, originally thought to have a dinuclear copper-active site, but which has very recently been suggested to be mononuclear, is housed on specialized intracytoplasmic membranes (17, 44), and can constitute a large proportion of total cellular protein. When copper levels are low, some methanotrophs (17, 45) have the ability to use the soluble MMO (sMMO), which has a dinuclear iron-active site (46). The switchover between these MMOs is copper-regulated, and more detail about this process and methanotroph classification and metabolism can be found in Ref. 17. Understanding how methanotrophs manage and use copper has immense environmental relevance due to methane being a highly potent greenhouse gas, and it is also essential for prospective biotechnological applications of these organisms and their MMOs (47-49).

The ability to utilize large amounts of copper results in methanotrophs having highly interesting copper-handling systems. This includes methanobactin (Mbn) (17, 50-52), which has been considered comparable to certain iron-binding siderophores (53), and thus termed a chalkophore (50). Mbn is a modified peptide (Fig. 1, A and B) that is part of a highly specialized copper-uptake system, secreted to sequester this metal ion under limiting conditions (17, 50). The mbnA gene, which codes for leader (cleaved) and core (modified) peptides, has been identified in an operon along with proteins either shown or suggested to be involved in modification reactions, apo-Mbn export, and re-incorporation of Cu(I)-Mbn (17, 51, 55-59). Related Mbn operons are present in some non-methanotrophic bacteria (17, 51, 59). Work in our lab has found that Mbn bind Cu(I) with affinities in the 10^{20} to 10^{21} M^{-1} range and have Cu(II) affinities that are ~ 6-10 orders of magnitude weaker (52, 54, 60). We suggested oxidation could assist removal of the metal ion in cells (52), although a conformational change at the N terminus of the peptide now appears to be the most likely mechanism to promote release (60). Having characterized a range of Mbn operons isolated from spent media in which methanotrophs were grown at low copper concentrations (52, 54, 60), including determination of their high-resolution crystal structures (Fig. 1, A and B) and analysis of the Cu(I)-Mbn uptake process (Fig. 1, C and D) (52), understanding the fate of internalized Cu(I)-Mbn became our next aim. To try to isolate intracellular Cu(I)-Mbn, soluble extracts from the model switchover methanotroph M. trichosporium OB3b were separated using anion-exchange, followed by size-exclusion, chromatography, and fractions analysed for metals. A number of copper-containing peaks were observed, but none contained Mbn.

Although Mbn was not found within cells in these metalloproteomic studies, the observation of soluble copper pools in M. trichosporium OB3b extracts, whose abundance increased at higher copper concentrations, is extremely interesting. The complex mixture present in the major copper-containing fraction was further purified to identify constituent copper-binding proteins (35). Copper
abundance in fractions matched the intensity of a band on an SDS polyacrylamide gel at ~12 kDa, which was purified to near homogeneity (Fig. 2A). The intensity profile of no other metal tested, including manganese, iron and zinc, corresponded with this band, identified by peptide mass fingerprinting as an uncharacterized conserved hypothetical protein possessing 13 Cys residues (Fig. 2B). The protein has a predicted twin-arginine translocase (Tat) signal peptide, suggesting it is folded prior to export from the cytosol (61), and cleavage is likely (62) after Ala-24 (Fig. 2B). Overexpressed Gly-1 to Ala-122 (no signal peptide) forms a tetramer of four-helix bundles (Fig. 2C), with all Cys residues pointing into the cores of the monomers (Fig. 2D) (35). The protein has no disulfide bonds due to the Cys residues all being found on α-helices in a fold that constrains the side chains. The protein can bind up to 12-14 Cu(I) ions per monomer in vitro, with an average Cu(I) affinity of ~1 × 10¹⁷ M⁻¹ (35).

In the crystal structure (35) each monomer binds 13 Cu(I) ions (Fig. 2E) in an unprecedented arrangement along the core of the four-helix bundle, a motif commonly found in metalloproteins (35, 63) including that binding the iron site of sMMO (46). Four of the Cu(I) ions are coordinated by two thiolates on the same α-helix in CXXXC motifs (Fig. 2B), whilst the majority of the other sites are ligated by two Cys residues on different helices. The solvent-accessible sites at the mouth of the bundle (the opposite end contains a number of hydrophobic side chains), via which Cu(I) ions are presumed to enter and leave, have different coordination environments (Fig. 2E) (35). This includes the binding of Cu13 by His-36, and Met-48 acting as a bridging ligand between Cu11 and Cu13. A tetramer capable of accommodating up to 52 largely solvent-protected Cu(I) ions is consistent with a role in storage, and hence the name of this novel family of copper proteins (the Csps) was devised (35). The way Csps bind Cu(I) is very different from the way a Cys-rich unstructured MT polypeptide folds around thiolate-coordinated clusters (Fig. 2F) (38, 39, 41). Ferritins, which store iron, are also four-helix bundles, but they use these to form a multimeric envelope that can be filled with thousands of Fe(III) ions (64). *M. trichosporium* OB3b possess three Csp homologues (Fig. 2G), and the first discovered as described above was called Csp1 (*Mt*Csp1 indicates it originates from *M. trichosporium* OB3b), and the others named *Mt*Csp2 and *Mt*Csp3.

**Comparison of Csp homologues in *M. trichosporium* OB3b**

*Mt*Csp1 and *Mt*Csp2 have high identity (~60% sequence conservation) and both possess predicted (61) Tat signal peptides. The presence of two exported Csps within the same methanotroph is not uncommon (see under “Csp homologues in other methanotrophs”), where they could perform different roles (see under “Functions of Csps”). *Mt*Csp3 has no signal peptide (Fig. 2G), is therefore cytosolic, and has lower sequence identity to *Mt*Csp1 and *Mt*Csp2 (~20% conserved residues). Neither *Mt*Csp2 nor *Mt*Csp3 have been identified by metalloproteomics, although many of the other soluble copper-containing fractions obtained from *M. trichosporium* OB3b have yet to be thoroughly investigated. *Mt*Csp2 has not been studied due to its high sequence similarity to *Mt*Csp1 (Fig. 2G). *In vitro* studies of *Mt*Csp3 show it is also a tetramer of four-helix bundles having 18 Cys residues pointing into the core of each monomer (Fig. 3A) (36). The additional Cys residues, compared to *Mt*Csp1, are found in CXXXC motifs, and the protein also has no disulfide bonds. Each monomer is able to bind more metal ions within its core and has 19 Cu(I) sites in the crystal structure (Fig. 3B). Most of these are coordinated by two thiolates, largely alternating between Cu(I) ions bound by Cys residues from the same α-helix (in CXXXC motifs) and inter-helical sites. Atypical coordination is again found at the mouth of the bundle where His-110 binds Cu18 along with Cys-111, and His-104 ligates Cu19 in addition to two thiolates (Fig. 3B) (36). The average Cu(I) affinity of *Mt*Csp3 (~2 × 10¹⁷ M⁻¹) is similar to that of *Mt*Csp1 (35, 36).

Important differences are found in how *Mt*Csp3 and *Mt*Csp1 bind Cu(I). In the case of *Mt*Csp3, Cu(I) binding gives rise to relatively intense fluorescence at ~600 nm upon excitation within the S(Cys) → Cu(I) ligand-to-metal charge transfer bands below 400 nm (35, 36). Such emission has been associated with Cu(I)-Cu(I) interactions in proteins binding solvent-protected Cys-coordinated Cu(I) clusters, such as the MTs (38, 42, 66). The fluorescence from *Mt*Csp3, which reaches a maximum value when it is approximately half-loaded, may be related to the formation of solvent-protected tetranuclear Cu(I) clusters within
its central core (67). Similar structures do not occur in \textit{MtCsp1} due to it having fewer Cys residues. Furthermore, and functionally more important, Cu(I) binding is cooperative in \textit{MtCsp1} (35), but not in \textit{MtCsp3} (36). This could also be related to discrete cluster formation in \textit{MtCsp3} (67), but the exact cause of both of these aspects of Cu(I) binding in the Csp3 requires further investigation.

The most striking difference between \textit{MtCsp3} and \textit{MtCsp1} is the time scale of Cu(I) removal from their cores. Both have average Cu(I) affinities in the low $10^{17}$ M$^{-1}$ range, and assuming diffusion-controlled on-rates of $\sim 10^8$ M$^{-1}$s$^{-1}$, unassisted Cu(I) off-rates would be extremely slow ($\sim 10^{-9}$ s$^{-1}$). The physiological Cu(I) acceptor for any unassisted Cu(I) off-rates would be extremely slow Cu(I) binding is cooperative in

Furthermore, and functionally more important, Cu(I)-affinity molecule (54, and OB3b produces an Mbn, removal by this high Cu(I)-chelating molecules such as bicinchoninic acid, and particularly bathocuproine disulfonate (BCS). These ligands have routinely been used to

Comparative Cu(I) removal studies have also been carried out with well-characterized chromophoric Cu(I)-chelating molecules such as bicinchoninic acid, and particularly bathocuproine disulfonate (BCS). These ligands have routinely been used to measure how tightly Cu(I) binds to a range of proteins (26, 68-71), including determination of the Cu(I) affinities of Mbns (52, 54) and the average values for the Csp3s (35, 36). They have also been implemented as model acceptors for investigating Cu(I) removal from homeostasis proteins (68, 72). Using a large excess of the higher-affinity ligand (BCS) results in complete removal of Cu(I) from MtCsp1 in $\sim$1 h (35), while this process takes weeks to complete for MtCsp3 (36). Comparative Cu(I) removal studies have also been carried out with well-characterized chromophoric Cu(I)-chelating molecules such as bicinchoninic acid, and particularly bathocuproine disulfonate (BCS).

Csp homologues in other methanotrophs

Homologues of \textit{MtCsp3} are present in 34 methanotrophs whose genomes have been sequenced, whilst \textit{MtCsp1} homologues are found in 16 (Fig. 4A and Figs. S1 and S2). A single \textit{MtCsp3} homologue is typically found in methanotrophs having this protein (two in \textit{Methylcococcaceae bacterium} NSP1-2 and \textit{Crenothrix polyspora}), and two \textit{MtCsp1} homologues are present in over half of the sequenced methanotrophs that have this protein (\textit{Methylcystis bryophila} appears to have three, although one of these has only seven Cys residues). A different name for a protein implies an alternative function. However, for organisms with two or more \textit{MtCsp1} homologues it is not yet known whether these have distinct functions (see below). We have not established a way to differentiate between what we initially called \textit{McCsp1} and \textit{McCsp2} (35), and it may therefore be better to use Csp1a (\textit{McCsp1a}) and Csp1b (\textit{McCsp1b}) to signify exported Csp homologues when found within the same organism. This approach provides the clear definition that Csp1s are the exported members of this family of proteins, whereas Csp3s are cytosolic. For the purposes of this review, we will continue to use \textit{MtCsp1} and \textit{MtCsp2} for the exported proteins in \textit{M. trichosporium} OB3b.

Bioinformatics also highlight residues and regions of Csp1s and Csp3s that are conserved in methanotrophs (Fig. 4, B and C and Figs. S1 and S2). This includes the Cu(I)-coordinating Cys residues; all 13 are highly conserved in the Csp1s (Fig. 4B), with 15 highly conserved in Csp3s (Fig. 4C). The sequence of the Tat signal peptide is highly similar in the Csp1s (Fig. 4B). Conserved regions are found at the open end of both four-helix bundles, thought to be important for Cu(I) uptake and removal. The His-36 and Met-48 ligands in \textit{MtCsp1} are present in almost all homologues, along with a number of the intervening residues in the \alpha1(His-36)-loop-\alpha2(Met-48) region. In addition, two residues (Phe-97 and Pro-98 in \textit{MtCsp1}) on the loop linking \alpha3 and \alpha4 are highly conserved in Csp1s, as is the Glu residue (Glu-102 in \textit{MtCsp1}) at the start of \alpha4 (Fig. 4B). Certain residues in the \alpha3-loop-\alpha4 region are conserved in Csp3s. This includes the His-110 ligand of \textit{MtCsp3} (Fig. 3B) and the non-ligating His-108, but these are predominantly present only in \textit{Methylcystis cystaceae} family (\textit{Methylasinus, Methylcystis} and \textit{Pleomorphomonas} genera) strains. His-104, which coordinates Cu18 in \textit{MtCsp3}, is also conserved in these strains, but is replaced by an Ile residue in
most other methanotrophs that have a MtCsp3 homologue (Fig. 4C and Fig. S2). Overall, the Cys ligands (Csp1s and Csp3s), the Tat signal peptide (Csp1s), the α1-loop-α2 (Csp1s) and α3-loop-α4 (Csp1s and certain Csp3s) regions are conserved features of these proteins in methanotrophs.

The Mbn operon is present in 12 methanotroph genomes, and due to overlap, less than half of sequenced methanotrophs (38 of 89) possess either the Mbn operon, Csp1 or Csp3 (Fig. 4A). Nine methanotrophs possess all three, with none having the Mbn operon without either a Csp1 or Csp3. Overlap mostly occurs in the Methylosinus and Methylocystis genera, but there is not sufficient evidence to suggest the functions of Csp1, Csp3 and Mbn are directly related. However, the observation that Csp3 is present alone in 22 methanotrophs (Csp1 is rarely present on its own) could indicate the function of this protein is independent of Csp1 and Mbn (see under “Functions of Csps”). Much attention has been paid to the role of Mbn in copper acquisition and utilization by methanotrophs (17, 50-52, 54-60). However, the currently identified Mbn operon is only found in just over 10% of sequenced methanotroph genomes, suggesting that Csp1, Csp3 and Mbn (see under “Functions of Csps”). Much attention has been paid to the role of Mbn in copper homeostasis and Mbn (Fig. 4B) (75). Crystal structures of the apo-protein, including its tetrameric arrangement, is similar to MtCsp3 (Fig. 3, A and C), again with little evidence of disulfide bond formation (36). The protein binds up to ~20 eq of Cu(I) in vitro with an average affinity of ~2 × 10^17 M^-1, but removal is faster than for MtCsp3 as ~85% of the Cu(I) core is acquired by BCS in 85 h (36). However, this is still very slow compared to MtCsp1 (35), and the mouth of the four-helix bundle of BsCsp3 shows similarities to that of MtCsp3 (Fig. 3, A and C) (36). In particular, three His residues corresponding to His-104, His-108 and His-110 are present, and these are more conserved in non-methanotrophs (36). The cytosolic Cu(I) metallochaperone CopZ is present in B. subtilis. Removal of Cu(I) by this potential physiological partner is also slow with BsCopZ acquiring ~40% of Cu(I) from BsCsp3 in 64 h (36). The in vitro Cu(I)-binding properties and structure of Streptomyces lividans Csp3 are similar to those of MtCsp3 and BsCsp3 (36), although the average Cu(I) affinity appears to be an order of magnitude weaker (75). Crystal structures of the apo-Csp3s from Pseudomonas aeruginosa (3KAW) and Nitrosospira multiformis (3LMF) have been deposited by a structural genomics consortium. These are similar to those of other Csp3s (Fig. 3, A and C), with all Cys residues pointing into the cores of their four-helix bundle folds, and no disulfide bonds present.

Functions of Csps

The predicted Tat-exported MtCsp1 acts as a copper store for methane oxidation for the following reasons. 1) copper-bound MtCsp1 is isolated (35) from M. trichosporium OB3b grown in 5 μM copper (Fig. 2A) i.e. using pMMO to oxidize methane. None of the other metals analyzed in the metalloproteomic studies (including
manganese, iron, and zinc) co-eluted with MtCsp1. Furthermore, of the metal ions tested in our laboratory, only Cu(I) binds tightly in vitro. 2) The deletion of both genes for the exported Csps (MtCsp1 and MtCsp2) results in significantly faster switchover from pMMO to sMMO in M. trichosporium OB3b cells transferred from high to low copper (35). 3) In gene expression studies (see Fig. S5 in Ref. 58 and see also Ref. 78) MtCsp1 is up-regulated in a similar manner to pMMO at a copper concentration resulting in switchover (10-12.5 μM). 4) The copper peak and the MtCsp1 SDS-PAGE band (Fig. 2A) are absent in sMMO-active M. trichosporium OB3b cells. 5) The structure of MtCsp1 as a tetramer of Cys-lined four-helix bundles allows the binding of 52 Cu(I) ions. Collectively, these data provide extensive evidence that MtCsp1 stores Cu(I) for pMMO allowing continued growth on methane using this enzyme when copper becomes limiting. Given the similarities of the structures and Cu(I)-binding properties of MtCsp1 homologues, many having an even greater capacity for metal ions, it is almost implicit that other members of this new family of proteins are able to bind and store Cu(I).

The cellular destination of exported Csp1s depends on the cellular structure of methanotrophs (see below). The periplasmic multi-copper oxidase CueO, which is involved in copper homeostasis in E. coli, is also predicted to be Tat-exported (79). As this is the pathway for folded protein secretion, it had been assumed that CueO acquired the four copper ions it needs for activity in the cytosol. However, it is now thought this protein is exported in a copper-free “incomplete folding” state and acquires copper in the periplasm (79). A number of other bacterial copper proteins are predicted to be Tat exported (7), and further work is needed to determine whether these acquire copper in the cytosol. However, it seems highly unlikely that Csp1s are exported in a partially folded state as this would potentially promote disulfide bond formation in such Cys-rich proteins, and Tat-export may be required to prevent this occurring. Csp1s therefore most likely fold completely, and acquire Cu(I), in the cytosol, prior to export.

Although the prevailing view is that the intracytoplasmic membranes housing pMMO are invaginations of the plasma membrane (Fig. 4D), most evidence is either out-dated or indirect (80-82). If these membranes are discrete from the plasma membrane (Fig. 4D), pMMO would be only the second example, after plastocyanin in the thylakoid compartments of cyanobacteria (83), of a bacterial cytoplasmic copper-requiring protein. If this is the case, in a methanotroph such as M. trichosporium OB3b having two exported Csps, MtCsp1 could deliver Cu(I) to the intracytoplasmic membranes for pMMO, while MtCsp2 transfers Cu(I) to the periplasm for other copper-requiring enzymes (Fig. 4D). MtCsp2 is not up-regulated by 10 μM copper (78), and this suggested function may not require copper-regulated expression (see below), or may occur at higher copper concentrations than those causing switchover (it is also possible that MtCsp2 could act as a Cu(I) store for pMMO at higher copper concentrations). In methanotrophs that have a single Csp1, we assume this will only store Cu(I) for pMMO if the enzyme is housed in cytoplasmic compartments. However, if the intracytoplasmic and plasma membranes are contiguous, a single Csp1 could store Cu(I) for pMMO and other destinations in the periplasm. As already stated, most Csp1-possessing methanotrophs also have Mbn (Fig. 4A), which they produce under copper-limiting conditions (17, 50), when the Cu(I) from a Csp would be required. Therefore, during switchover from pMMO to sMMO it is possible that apo-Mbn may play a role in removing Cu(I) from MtCsp1, a process that readily occurs in vitro (35), to aid delivery to pMMO. In non-methanotrophs, Csp1 will deliver Cu(I) to the periplasm. Export of a protein that can store large amounts of Cu(I) (there is little sign of oxidation upon prolonged exposure of Cu(I)-MtCsp1 in air) will provide and stabilize a source of cuprous ions outside the cytosol, which may otherwise be difficult in the more oxidizing periplasm. This could be the oxidation state of copper required for insertion into certain enzymes, as appears to be the case for pMMO.

The precise function of a cytosolic Csp3 is currently unknown, although a general role in Cu(I) storage while preventing toxicity is presumed. Bioinformatics (Fig. 4A), and the more widespread nature of Csp3s both in methanotrophs and other bacteria, suggest that the function of Csp3 is not directly linked to Csp1. Preliminary in vivo studies on the csp2-delete strain of B. subtilis show a weak, and unusual, copper-dependent phenotype (36). Growth in LB media is inhibited relative to WT B. subtilis in the range of ~1.5 to 2 mM added Cu(I),
but only after cells have been grown for more than 12 h in the presence of the metal. Obtaining this phenotype reproducibly is difficult, being sensitive to growth conditions, and particularly copper concentration. Transcriptional studies have shown that BsCsp3 is up-regulated under spore-forming and stress-inducing conditions, including elevated NaCl concentrations (84), but the response to copper was not tested. Interestingly, both Csp1 and Csp3 are up-regulated when the methanotroph Methylocystis sp. SC2 is grown in 0.75% NaCl (85). The relevance of salt stress on Csp expression remains unclear. The multi-copper oxidase CotA is one of very few predicted copper enzymes in B. subtilis (2) and is a component of the spore coat where it is thought to be involved in pigment production (86), and BsCsp3 could store Cu(I) for this enzyme. The P. aeruginosa Csp3 (the protein was incorrectly called a Csp1, but it does not possess a signal peptide and is cytosolic) is not induced (87) by the addition of 0.5 mM Cu(II). However, neither is CopA2, a second copper efflux pump that is not required for copper tolerance in P. aeruginosa, but is suggested to be involved in export coupled to copper acquisition by cytochrome $c$ oxidase (88). The S. lividans Csp3 is up-regulated by 0.4 mM Cu(II) and in a csoR (copper-sensitive operon repressor) deletion mutant (89). A transcriptomic study of the Gram-negative bacterium Sphingobium sp. ba1 has shown up-regulation of Csp1 and Csp3 in response to 10 mM Ni(II), but under these conditions copper resistance systems, including CopA, are also up-regulated (90).

Gene expression studies show that MtCsp3 is not up-regulated at the relatively low copper concentrations (10 μM) required for switchover (78). However, putative CopAs in M. trichosporium OB3b are also not up-regulated under these conditions. Copper detoxification is not the proposed primary function of Csp3s, but BsCsp3 can provide protection against copper toxicity when overexpressed in both the copA delete strain ($\Delta$copA) (Fig. 3D) (36) and also WT E. coli (Fig. 3E). In both cases, the cells overexpressing BsCsp3 accumulate more copper than control cells, and Cu(I)-BsCsp3 is observed. As well as being able to complement the phenotype caused by deletion of the copper-efflux pump (23), overexpressed BsCsp3 provides an additional growth advantage at elevated copper to having CopA alone. Furthermore, BsCsp3-bound Cu(I) can be withheld from the efflux pump. It has also been found that in S. lividans Csp3 enables growth at higher copper levels (75).

When considering the functional properties of Csp3s it is important to keep in mind key in vitro results (35, 36). Csp3s can generally bind a greater number of Cu(I) ions than Csp1s, due to usually having more Cys residues (Fig. 2, E and G, Fig. 3B, Fig. 4, B and C, and Figs. S1 and S2 for methanotrophs). Csp1s and Csp3s have similar average Cu(I) affinities (~$10^{17}$ M⁻¹) yet exhibit dramatic differences in terms of Cu(I) removal rates. How Cu(I) is extracted from Csp3s in cells is unknown. Many copper homeostasis proteins and copper target enzymes/proteins have higher Cu(I) affinities, typically in the $10^{17}$ to $10^{21}$ M⁻¹ range (13, 26, 40, 68-71), and their ability to acquire Cu(I) from Csp3s is thermodynamically favored. Faster Cu(I) unloading by small molecule Cu(I) ligands occurs for Csp1s (35), but a kinetic barrier to removal is present in Csp3s (36). The interplay between thermodynamics and kinetics in copper homeostasis is currently not well understood. Furthermore, how many of the proteins involved in this process acquire copper is unknown (apart from the CopZ/CopA interaction). Csp1s are expected to be exported after acquiring Cu(I), and Csp3s kinetically trap Cu(I) in the cytosol. These proteins may therefore have evolved different approaches to enable them to bind and maintain a store of Cu(I) even in the presence of proteins with higher affinities (CopZ acquires Cu(I) very slowly from BsCsp3). In methanotrophs, which can have both a Csp1 and Csp3, such as M. trichosporium OB3b, the dramatic variation in removal rates could be more important (as may differences in Cu(I)-binding cooperativity), and suggests that the exported Csp3s act as a more temporary store of Cu(I) for pMMO, whilst CopA plays a role in longer-term storage. Whether this distinction between Cu(I) removal rates exists for all Csp1s and Csp3s has to be established. If slow Cu(I) removal is a conserved feature of Csp3s, then the requirement for a longer-term store needs to be understood, as well as how the kinetic barrier to removal is overcome when Csp3-bound Cu(I) is required.
Possible link between Csps and pathogenicity?

As mentioned in the Introduction, and covered in previous minireviews in this series, the interplay between copper homeostasis systems in a pathogen and host is beginning to be recognized as important for virulence (30-34, 91). Compared to nutritional immunity used to withhold other essential metal ions, hosts are thought to expose invading pathogens to copper (32, 33, 91). In mammalian hosts, ATP7A pumps copper into the phagolysosomal compartment, and copper homeostasis systems can protect the pathogen against this attack (30-33, 91-93). A number of possible defence approaches have been identified, such as copper efflux and sequestration, including by a Cu(II)-binding siderophore (34, 94). Csps are present in pathogenic bacteria, such as N. gonorrhoeae (Csp1), Streptococcus pneumoniae (Csp3), Salmonella enterica sv. Typhimurium (Csp3), and the opportunistic pathogen P. aeruginosa (Csp3). The ability of Csps to bind large quantities of Cu(I) would make them ideal to defend pathogens against copper attack by a host. The Cu(I)-buffering ability of Csp3 that prevents toxicity in the ΔcopA strain of E. coli (Fig. 3, D and E) demonstrates that these proteins, when produced at relatively high levels, can take the place of copper-efflux pumps, known virulence factors (30, 88, 92, 93), in providing protection against elevated copper levels. The only other characterized bacterial copper-storing protein is the MT-like MymT found in pathogenic mycobacteria (42), but this does not appear to be required for infection. Whether a Csp would help a pathogen fight against host-based copper attack remains to be established.

Concluding remarks

The Csps were identified in methanotrophs, bacteria with atypically high copper demands, which use it for methane oxidation, and M. trichosporium OB3b possesses three homologues: two closely related proteins having predicted Tat signal peptides and a cytosolic version. Exported McCsp1 stores Cu(I) for pMMO. The more widespread occurrence of cytosolic Csps, complicates the current conceptually simplistic idea that these organisms have evolved not to use copper in this compartment to help avoid toxicity. A role for these proteins in Cu(I) storage is currently the most logical suggestion for their function, but in many cases what they are storing copper for remains unknown. The presence of bacterial copper storage proteins seems consistent with a number of other observations as follows: 1) that bacterial copper-import systems exist (6, 7, 17, 21, 52, 56, 77, 95), including into the cytosol; 2) that endogenous pools of the metal are available in bacteria (11, 15, 16, 18, 96); and 3) that E. coli grown in both LB and minimal medium accumulates copper (97). It also suggests that there are alternative mechanisms to using different cellular compartments to prevent mis-metallation of proteins by copper (37). Furthermore, the ability of bacteria to store copper in the cytosol could provide further insight into the observation that certain periplasmic proteins are loaded with copper that has passed through the cytosol (88, 98, 99). Most of the organisms in which this has been reported possess a Csp3.

A lot more work is needed to understand copper storage and removal for the exported Csp1s and the cytosolic Csp3s. Csps are only found in ~40% of methanotrophs, whilst pMMO is nearly always present, and although Csp3s are widespread in bacteria they are far from ubiquitous. The discovery of the Csps leads to the intriguing question of whether there are other bacterial copper-storage systems yet to be found. Even if this is not the case, the presence of Csp3s indicates that as predicted for other metalloproteomes (100), the possibility exists that there are cytoplasmic copper-requiring enzymes yet to be discovered.

Acknowledgments: We thank everyone who has worked on the Csp project to date including (in alphabetical order); Stephen Allen, Arnaud Baslé, Owen Burbidge, Andrew Crombie, Claudia Fogl, Gianpiero Landolfi, Colin Murrell, Neil Patterson, Semeli Platsaki, Nicolas Vita and Kevin Waldron.

Conflict of interest: The authors declare that they have no conflicts of interest with the contents of this article.
References


The three-dimensional structure of the cyanobacterium *Synechocystis* sp. PCC 6803. *Arch. Microbiol.* **184**, 259–270


FOOTNOTES
This work was supported by the Natural Environment Research Council Grant NE/F00608X/1, the Biotechnology and Biological Sciences Research Council Grant BB/K008439/1, a studentship (to S. D.), and by Newcastle University. This is the second article in the Thematic Minireview series “Metals in Biology 2018: Copper homeostasis and utilization in redox enzymes.”

This article contains Figs. S1 and S2 and supporting Refs. 1–3.

The abbreviations used are: ROS, reactive oxygen species; MT, metallothionein; Csp, copper storage protein, pMMO, particulate methane monooxygenase; sMMO, soluble methane monooxygenase; Mbn, methanobactin; Tat, twin-arginine translocase; BCS, bathocuproine disulfonate; PDB, Protein Data Bank.
Figure 1. Structures of Cu(I)-Mbns and Mbn-mediated copper uptake. A and B, the crystal structures of the Cu(I)-Mbns from *M. trichosporium* OB3b (A, PDB file 2XJH) (54) and *Methylocystis hirsuta* CSC1 (B, PDB file 2YGI) (52). Below the structures are the sequences of the leader (black and underlined) and core peptides that make up MbnA. Core peptides are modified to give the Mbn, and the *M. hirsuta* CSC1 Cu(I)-Mbn structure is of a form with the three C-terminal residues cleaved (amino acids are numbered according to the sequence of the core peptides). The Cu(I) ions are shown as orange spheres ligated by the sulfur atoms (S¹ and S²) from thioamide/enethiol groups, and two oxazolone (oxa) ring nitrogens in *M. trichosporium* OB3b Cu(I)-Mbn, with the N-terminal coordinating heterocycle being a pyrazinediol (pyra) in *M. hirsuta* CSC1 Cu(I)-Mbn. Other differences include a sulfate-modified Thr side chain in *M. hirsuta* CSC1 Cu(I)-Mbn and the overall hairpin-like structure of this Cu(I)-Mbn compared to the more compact *M. trichosporium* OB3b Cu(I)-Mbn. Also shown are copper uptake by (lines) and relative sMMO activity of (bars) *M. trichosporium* OB3b (C) and *M. hirsuta* CSC1 (D) cells after the addition of *M. trichosporium* OB3b Cu(I)-Mbn (open grey triangles and grey bars) and *M. hirsuta* CSC1 Cu(I)-Mbn (open cyan circles and cyan bars) to sMMO-active cells. In both cases, copper uptake and switchover from sMMO to pMMO is faster with the native Cu(I)-Mbn (52).
Figure 2. Discovery and characterization of Csp1 in M. trichosporium OB3b. A, copper content of anion-exchange soluble extracts of M. trichosporium OB3b purified on a gel-filtration column, and SDS-
PAGE analysis of the fractions eluted between 8 and 15 ml (35). The intensity profile of the band indicated with an arrow matches that of the main copper peak. B. Sequence of this 146 amino acid residue protein (MtCsp1) has a predicted Tat signal peptide (bold) and 13 Cys residues (highlighted yellow) largely present in CXXXC and CXXC motifs (underlined). C. Tetrameric arrangement in the crystal structure (PDB file 5FJD) (35) of the overexpressed predicted mature form of MtCsp1 (Gly-1 to Ala-122) with the pink (top left) four-helix bundle monomer (helices numbered), shown in detail (D), highlighting the Cys residues that all point into the core, and residues around the mouth of the bundle, in stick representation. E. Structure (PDB file 5FJE) of Cu(I)-MtCsp1 (35) with the metal ions represented as grey spheres and numbered. F. Crystal structure (PDB file 1RJU) of a truncated form of Saccharomyces cerevisiae MT binding eight Cu(I) ions via 10 Cys residues (39) is shown for comparison. G. Sequence alignment of the three Csp homologues present in M. trichosporium OB3b created in T-coffee (65) using the predicted mature forms of MtCsp1 and MtCsp2.
Figure 3. Structural and functional studies of Csp3s. A, crystal structure of the apo-MtCsp3 (PDB file 5ARM) tetramer (36). The side chains of the 18 Cys residues pointing into the core and three His residues at the mouth are shown as sticks in the pink (top left) four-helix bundle monomer. The additional small N-terminal α-helix (αN) is labeled in two monomers. B, crystal structure (PDB file 5ARN (36)) of Cu(I)-MtCsp3 (αN omitted) with the metal ions as grey spheres and numbered. C, crystal structure of the apo-BsCsp3 (PDB file 5FIG) tetramer of four-helix bundles (36) using the same representation as in (A), with 19 Cys residues pointing into the core of the pink (top left) four-helix bundle monomer. D, growth (37 °C) of ΔcopA E. coli in the absence of (black circles) and plus 1.0 mM (blue circles) Cu(NO₃)₂ is compared with ΔcopA cells overexpressing BsCsp3 in the absence of (red triangles) and plus 1.0 mM (cyan triangles) Cu(NO₃)₂ (36). E, growth (37 °C) of WT E. coli in the absence of (black circles) and plus 3.4 mM (blue circles) Cu(NO₃)₂ is compared with WT cells overexpressing BsCsp3 in the absence of (red triangles) and
plus 3.4 mM (cyan triangles) Cu(NO$_3$)$_2$. Overexpressed $\text{BsCsp3}$ can protect $\Delta\text{copA (D)}$ and WT (E) $\text{E. coli}$ from copper toxicity.
Figure 4. Bioinformatics of Csp1s, Csp3s and the Mbn operon in methanotrophs. The 89 methanotroph genomes currently available in the NCBI database were interrogated with pBLASTp using MtCsp1 and MtCsp3 as search queries. A, Venn diagram of the distribution of Csp1, Csp3 and the Mbn operon (identified by the presence of homologues of _M. trichosporium_ OB3b MbnA, MbnB and MbnC) in methanotroph genomes. Alignments of 26 and 36 sequences (see Figs. S1 and S2) were used to produce WebLogos (73) for Csp1s (B) and Csp3s (C), respectively. In these, the overall height of the stack at a particular position represents the degree of conservation, whereas the height of the symbol for an amino acid residue (green for polar, purple for neutral, blue for basic, red for acidic and black for hydrophobic) within the stack signifies relative frequency. Widths are unscaled so less frequently occupied positions in the alignment (see Figs. S1 and S2) are not represented by narrower stacks (composition adjustment was left to the default value for typical amino acid usage). Signal peptides were identified using SignalP (74) and TatP (62), and this region is labelled in B. D, schematic of the model methanotroph _M. trichosporium_ OB3b showing two possible arrangements of the intracytoplasmic membranes that house pMMO, and the location and potential roles of MtCsp1, MtCsp2 and MtCsp3. The established cytosolic copper-sensing (CueR) and copper-efflux (CopA) systems, and the known locations and interactions of Mbn (MbnT is a Ton-B-dependent transporter (56) and MbnE is suggested to bind Mbn in the periplasm (58)) are also included. A much more detailed model of copper homeostasis and copper-regulated switchover in a methanotroph can be found in Ref. 59. This includes the Csp, but does not discuss their importance.
SUPPORTING INFORMATION

Bacterial copper storage proteins

Christopher Dennison,* Sholto David and Jaeick Lee

From the Institute for Cell and Molecular Biosciences, Medical School, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK
Methylosinus trichosphorium(1)  GTNSAFTPAFVAVDVCACKKECKFKFSSIAECKKCAAECKK  119
Methylosinus trichosphorium(2)  AVNSSATPAALKTVYDVCMACKKECDRFPQYSECKNCACKQAK  122
Methylomagnum  ------------------------------------------GECRENKVDNNGAAAYQAYKEKKE  57
Methylcystis SC2  AANSSHVGPLAKTVAMICEDCKKQCDKFPKIAECKKCAAECKK  124
Methylcystis LW5(1)  AVNSTPTPLAKAVGDICMACKKECDRFPQYSECKACGDACKACAEECRK  121
Methylcystis LW5(2)  GVGSALTPAFAKVVAEACLACKKECDKFPNIAECKACGDACKACADDCQK  118
Methylcystis SB2  ALNSSHVGPLAKTVAMICEDCKKQCDKFPKIAECKKCAAECKK  124
Methylcystis bryophila(1)  AVNSSPTPLAKAVGDICMACKKECDRFPQYSECKACGDACKACADDCQK  118
Methylcystis bryophila(3)  ASDAPFAYGFARTAADLCAVSRRDCEKFPKIAECAALAAAAACEAACRK  113
Methylcystis rosea  AVNSSHVGPLAKTVAMICDDCKKQCDKFPKVKECVECGKSCQKCADECRK  123
Methylosinus LW4(1)  GTASALTPSFAKAVQAECFACKKECDKFPNIAECKACGDACKACADDCQK  118
Methylosinus LW4(2)  AVNSTPTPLAKAVGDICMACKKECDRFPQYSECTACGDSCKACADDCQK  121
Methylosinus PW1(1)  GVGSALTPAFAKVVAEACLACKKECDKFPNIAECKACGDACKACADDCQK  118
Methylosinus PW1(2)  AVNSTPTPLAKAVGDICMACKKECDRFPQYSECKACGDACKACAEECRK  121
Methylosinus R-45379(1)  AVNSTPTPLAKAVGDICMACKKECDRFPQYSECKACGDACKACADDCQK  118
Methylosinus R-45379(2)  GVGSALTPAFAKVVAEACFACKKECDKFPNIAECKACGDACKACADDCQK  118
Methylosinus LW3(1)  AVNSTPTPLAKAVGDICMACKKECDRFPQYSECKACGDACKACADDCQK  121
Methylosinus LW3(2)  GVGSALTPAFAKVVAEACFACKKECDKFPNIAECKACGDACKACADDCQK  118
Methylosinus 3S-1(1)  GTNSAFTPAFVAVDVCACKKECKFKFSSIAECKKCAAECKK  119
Methylosinus 3S-1(2)  AVNSSATPAALKTVYDVCMACKKECDRFPQYSECKNCACKQAK  122
Methylocapsa palsarum  AVNSPHTGHFAKVVEMICTDCQKECEKFPKIAECKKCAAECKK  116
Methylocapsa acidiphila(1)  SVNSPHTGHFAKVVEMICSDCQKECEKFPKIAECKKCAAECKK  116
Methylocapsa acidiphila(2)  AVNSVHTGHFAKVVEMICNDCAKECEKFPHIAECKACGESCKKCAAECKK  116
Methylocapsa aurea  AVNSEHTGHFAKVVEMICADCQKECEKFPKIAECKKCAAECKK  116
Methylocystis ATCC 49242  STNSNFTPAMAKVVAGVCEACKKECDKFPEVAECNAMGAACKACADECKK  117

*::          . :.      *..
----------------------------------------------
α3  ----------------------  α4

Figure S1. Multiple sequence alignment of Csp1s from methanotrophs produced using T-coffee (1). Cysteine residues are highlighted yellow, signal peptides were predicted using SignalP (2) and TatP (3), and these are underlined in red. The α-helices (numbered) that make up the four-helix bundle of Methylosinus trichosphorium OB3b Csp1 are indicated with coloured lines. Asterisks indicate fully conserved sequence positions; the ‘:’ and ‘.’ symbols indicate strongly and weakly similar sequence positions respectively. The sequences used are (organism names are as found on the NCBI website using the listed protein accession code): Methylosinus trichosphorium(1) - Methylosinus trichosphorium OB3b, accession WP_003609243.1; Methylosinus trichosphorium(2) - Methylosinus trichosphorium OB3b, accession WP_003610915.1;
Methylomagnun – *Methylomagnun ishizawai*, accession WP_085215912.1; Methylocystis SC2 - *Methylocystis* sp. SC2, accession WP_014889938.1; Methylocystis LW5(1) - *Methylocystis* sp. LW5, accession WP_026598754.1; Methylocystis LW5(2) - *Methylocystis* sp. LW5, accession WP_026599776.1; Methylocystis SB2 - *Methylocystis* sp. SB2, accession WP_029651346.1; Methylocystis bryophila(1) - *Methylocystis bryophila*, accession WP_085770723.1; Methylocystis bryophila(2) - *Methylocystis bryophila*, accession WP_085773360.1; Methylocystis bryophila(3) - *Methylocystis bryophila*, accession WP_085770896.1; Methylocystis rosea - *Methylocystis rosea* SV96, accession WP_018408544.1; Methylosinus LW4(1) - *Methylosinus* sp. LW4, accession WP_026191566.1; Methylosinus LW4(2) - *Methylosinus* sp. LW4, accession WP_018265259.1; Methylosinus PW1(1) - *Methylosinus* sp. PW1, accession WP_036293589.1; Methylosinus PW1(2) - *Methylosinus* sp. PW1, accession WP_036286969.1; Methylosinus R-45379(1) - *Methylosinus* sp. R-45379, accession WP_064032970.1; Methylosinus R-45379(2) - *Methylosinus* sp. R-45379, accession WP_026599776.1; Methylosinus LW3(1) – *Methylosinus* sp. LW3, accession WP_024880072.1; Methylosinus LW3(2) – *Methylosinus* sp. LW3, accession WP_024878120.1; Methylosinus 3S-1(1) - *Methylosinus* sp. 3S-1, accession WP_003609243.1; Methylosinus 3S-1(2) - *Methylosinus* sp. 3S-1, accession WP_003610915.1; Methylocapsa palsarum - *Methylocapsa palsarum*, accession WP_091681455.1; Methylocapsa acidiphila(1) - *Methylocapsa acidiphila* B2, accession WP_026607374.1; Methylocapsa acidiphila(2) - *Methylocapsa acidiphila* B2, accession WP_026605589.1; Methylocapsa aurea - *Methylocapsa aurea*, accession WP_036262121.1; Methylocystis ATCC 49242 - *Methylocystis* sp. ATCC 49242, accession WP_036279109.1. It is important to note that the two Csp1 sequences from *M. trichosporium* OB3b are identical to those of *Methylosinus* 3S-1. There are merely two differences between the *Methylosinus* sp. LW3(1), *Methylocystis* sp. LW5(1) and *Methylosinus* sp. R-45379(1) sequences, and only a single difference between the second homologue from these organisms (the second homologues from *Methylocystis* sp. LW5 and *Methylosinus* sp. R-45379 are identical). There are three differences between the sequence of *Methylosinus* sp. LW4(2) and *Methylosinus* sp. PW1(2), with the latter only having three differences to the *Methylosinus* sp. LW3(1) sequence. This very high sequence conservation is also true for MbnAs. The sequences from *M. trichosporium* OB3b and *Methylosinus* 3S-1 are identical, as are those from *Methylosinus* sp. PW1, *Methylocystis* sp. LW3 and *Methylosinus* sp. R-45379, with the MbnA from *Methylocystis* sp. LW5 having a single difference in its sequence.
Methylohalobius 10Ki                -------------DECIRLCLECAIACEVCL-YKM---SGLKSHNDCPVC  36
Methylocapsa                        ------ASDIGSLVRCVEACFDCAQACAACA-DACLGEDMVADLRSCIRL  59
Pleomorphomonas oryzae              ------GHIDAALLRCVEECLDCAQTCTSCA-DACLGESNVGDLRQCIRL  61
Pleomorphomonas koreensis           ------GTVDNDLLHCIEACLDCAQTCTSCA-DACLGEQNVDDLRQCIRL  61
Methylosinus LW3                    ------GRADAALVGCVEACFDCAQACAVCA-DACLTEENVADLRRCIRL  57
Methylosinus 3S-1                   ------GQTDRSLVQCVEMCFDCAQTCAACA-DACLGEDKVADLRHCIRL  57
Methylosinus R-45379                ------GRADAALVGCVEACFDCAQACAVCA-DACLTEENVADLRRCIRL  57
Methylosinus PW1                    ------GRADAALVGCVEACFDCAQACAVCA-DACLAEENVADLRRCIRL  65
Methylosinus LW4                    ------GRADAALVGCVEACFDCAQACAVCA-DACLAEENVADLRRCIRL  65
Methylocystis bryophila             ------GAENEALARCIEKCFDCAETCFACA-GACLDEEELKGLRECIRL  57
Methylocystis LW5                   ------GRADAALVGCVEACFDCAQACAVCA-DACLTEENVADLRRCIRL  57
Methylocystis SC2                   ---------PKEMQACIDACLNCYQMCFGLAMTHCLEKGGEHVKPKHFRA  45
Methylocystis ATCC 49242            ---------SKEMQSCVDECLRCYQMCFGMAMTHCLETGGDHVKPKHFRA  45
Crenothrix(1)                       --------HKLSLQACIAACNHCHQICLQTALNHCLEAGGKHVEATHFRL  48
Methylococcaceae NSP1-2(2)          ------FRTEQTIQAFVAACNLCHQLCLQAVVNY--KTGEKPIETDHLRL  59
Methylococcaceae NSP1-2(1)          --------TEHAMQDCIQACSHCHQVCLQTAMNHCLETGGKHVEAEHFRL  48
Methylovulum psychrotolerans        --------SVEAMQACITACCHCHKTCLQTAMNHCLETGGKHVAEEHFRL  48
Methylosarcina lacus                ------TPSEHATQACIEACSHCHQVCLATAMNHCLEAGGKHVKPKHFRL  50
Methylosarcina fibrata              ------TPSYQSMQSCIDICNRCAQICLQTAMNQCLEMGGRHVQPEHFRL  76
Methylomonas LWB                    --------SEQAMQSCLEACELCHRICLQTAMNHCLENGGKHVKAKHIRL  48
Methylomicrobium buryatense         --------TDSAMQACIETCTRCHQVCLQTAMNHCLETGGKHVEAEHFRL  48
Methylomicrobium alcaliphilum       --------TDSAMQACIETCTRCHQVCLQTAMNHCLETGGKHVEAKHFRL  48
Methylomicrobium album              --------TDSAMQACIETCTRCHQVCLQTAMNQCLEMGGRHVEPEHFRL  38
Methylomicrobium agile              --------MEEATRACIDACAHCHQTCLETAMNHCLETGGEHVAPEHFRL  49
Methylocaldum 14B                   --------STGSMQACIEQCLDCHSICLETV-THCLEKGGPHAEADHIRL  48
Methylobacter tundripaludum         --------TEHAMQACIEACSHCHQVCLQSAMNHCLKTGGKHVEAEHFRL  48
Methylobacter luteus                ------------MKSCIEMCNRCAGVCWKTAMNQCLEMGGRHVEPEHFRL  38
Methylobacter                       ------------MQSCIEMCTRCSQVCRETAMNQCLEMGGQHVEPEHFRL  38
----------------
Methylocapsa                        MV------MHIQ------------------AMIDTHP--S-----AR---  16
Pleomorphomas oryzae                MEAA----MHAE------------------AMIRSHP--A-----LN---  18
Pleomorphomas koreensis             MEAA----MHAE------------------AMIRSHP--A-----LN---  18
Methylocystis parvus                MHMM----------------------------------------------   4
Methylocystis bryophila             MH--------LQ------------------SMIAARP--K-----LQ---  14
Methylocystis LW5                   MH--------VE------------------AIIASHP--Q-----MR---  14
Methylocystis ATCC 49242            MHKM----------------------------------------------   4
Methylocystis (2)                   MNHS--------------------------T-----------------H-   6
Methylocystis (1)                   MNHS--------------------------A-----------------P-   6
Methylococcaceae NSP1-2(2)          MNYL--------------------------AN---AE--QAMKVWNY-L-  17
Methylococcaceae NSP1-2(1)          MNHS--------------------------T-----------------H-   6
Methylovulum psychrotolerans        MNPV--------------------------A-----------------H-   6
Methylovulum miyakonense            MDHL--------------------------S-----------------Q-   6
Methylosarcina lacus                MTQL--------------------------A-----------------H-   6
Methylosarcina fibrata              MKIAR------FYRSRLFILILENSTMYTSQS---EQ--K-----KT-S-  32
Methylomonas LWB                    MHQS--------------------------S-----------------N-   6
Methylocystis (2)                   MNHS--------------------------T-----------------H-   6
Methylocystis bryophila             MH--------LQ------------------SMIAARP--K-----LQ---  14
Methylohalobius 10Ki                ---QCAEHTHH-DHCLKCAESCRKCADACRQMA-------A 104
Methylocapsa                        ---ECRRHAQHHEHCRVCADSCADCEKACSIALQA--EVVH 137
Pleomorphomonas oryzae              ---ECGRHASMHEHCRVCADACRRCEAACAAAADGIRPSLQ 141
Pleomorphomonas koreensis           ---ECGRHATMHEHCRVCADACRHCEDACATAADGIKSSLQ 141
Methylosinus LW3                    ---ECRRHAKHHEHCRICADVCKECESACRRAMG----PAH 133
Methylosinus 3S-1                   ---ECRRHAGNHEHCRICADVCKECETACRSATG----LTH 133
Methylosinus R-45379                ---ECRRHAKHHEHCRICADVCKECESACRRAMG----PAH 133
Methylosinus PW1                    ---ECRRHAKHHEHCRICADVCKECESACRRAMG----PAH 141
Methylosinus LW4                    ---ECRRHAKHHEHCRICADVCKECESACRRAMG----PAH 141
Methylocystis parvus                ---ECDPIPD----MKDCAAECRRCAEECRKMSGR--KMAA 115
Methylocystis bryophila             ---ECESHAGHHEHCRICAEVCRACEAACNEAAP----GSH 133
Methylocystis LW5                   ---ECRRHAKHHEHCRICADVCKECESACRRAMG----PAH 133
Methylocystis SC2                   ---ECDPIPD----MKECADECRRCAEECRRMSGQ--KMAA 115
Methylocystis ATCC 49242            ---ECDALPD----MKDCAAQCRRCAEACRKMAGQ--KMAA 115
Crenothrix(2)                       ---DCEKIGG----MDECVAACRACAKSCRQMASV----QH 117
Crenothrix(1)                       ---SCERIGG----MNECVTACRDCATSCQQMVNI----QH 127
Methylococcaceae NSP1-2(2)          ---SCEKIGG----MNECVTACRDCATSCQQMVNI----QH 127
Methylococcaceae NSP1-2(1)          ---MCEKIGG----MNECVTACRDCATSCQQMVNI----QH 127
Methylovulum psychrotolerans        ---MCEKIGG----MNECVTACRDCATSCQQMVNI----QH 127
Methylovulum miyakonense            ---MCEKIGG----MNECVTACRDCATSCQQMVNI----QH 127
Pleomorphomonas koreensis           ---NLDCAEICQTSANFQLSS----SDFHRHLYGICADVCEMSAT--------  97
Crenothrix(1)                       ---DCEKIGG----MDECVAACRACAKSCRQMASV----QH 117
Methylocaldum 14B                   ---SCEQFEN-DPQMQACAQTCRSCADACQAMAEE--AGVE 121
Methylobacter trichosphorium         ---ECRRHAGNHEHCRICADVCKECETACRSATG----LTH 133
Methylobacter luteus                ---DCESIGD----MEECASTCRECAESCRQMASM---AMH 107
Methylocystis bryophila             ---NLCAEICQTSANFQLSS----SDFHRHLYGICADVCEMSAT--------  97
Methylocaldum 14B                   ---DCEKIGG----MDECVAACRACAKSCRQMASV----QH 117

\[\text{Methylohalobius 10Ki} \quad \text{Methylocapsa} \quad \text{Pleomorphomonas oryzae} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylovulum psychrotolerans} \quad \text{Pleomorphomonas koreensis} \quad \text{Crenothrix(2)} \quad \text{Methylocycloceae NSP1-2(1)} \quad \text{Methylocycloceae NSP1-2(2)} \quad \text{Methylocycloceae NSP1-2(1)} \"
Figure S2. Multiple sequence alignment of Csp3s from methanotrophs produced using T-coffee (1). Cysteine residues are highlighted yellow and the α-helices (numbered) that make up the four-helix bundle of Methylosinus trichosporium OB3b Csp3 are indicated with coloured lines. Asterisks indicate fully conserved sequence positions; the ‘:’ and ‘.’ symbols indicate strongly and weakly similar sequence positions respectively. The sequences used are (organism names are as found on the NCBI website using the listed protein accession code): Methylosinus trichosphorium - Methylosinus trichosporium OB3b, accession WP_003608458.1; Methylobacter - Methylobacter sp. BBA5.1, accession WP_020161380.1; Methylobacter luteus – Methylobacter luteus IMV-B-3098, accession WP_081679060.1; Methylobacter tundripaludum - Methylobacter tundripaludum SV96, accession WP_006889747.1; Methylobacter whittenburyi - Methylobacter whittenburyi, accession WP_036294836.1; Methylocaldum 14B - Methylocaldum sp. 14B, accession WP_077729496.1; Methylovolvum - Methylovolvum ishizawai, accession WP_085216804.1; Methylocaldum szegediense - Methylocaldum szegediense O-12, accession WP_026611966.1; Methylomicronium agile - Methylomicronium agile, accession WP_005369630.1; Methylomicronium album - Microcomirum album BG8, accession WP_031430794.1; Methylomicronium alcaliphilum - Methylomicronium alcaliphilum 20Z, accession WP_084685446.1; Methylomicronium buryatense - Methylomicronium buryatense 5G, accession WP_083877583.1; Methylomonas LWB – Methylomonas sp. LWB, accession WP_071157515.1; Methylomonas koyamae - Methylomonas koyamae, accession WP_064031195.1; Methylosarcina fibrate - Methylosarcina fibrate AML-C10, accession WP_020565470.1; Methylosarcina lacus - Methylosarcina lacus LW14, accession WP_024297779.1; Methylvolvulium miyakonense - Methylvolvulium miyakonense HT12, accession WP_019867731.1; Methylvolvulium psychrotolerans - Methylvolvulium psychrotolerans, accession WP_088620197.1; Methyloccaceae NSP1-2(1) - Methyloccaceae bacterium NSP1-2, accession OYV21462.1; Methyloccaceae NSP1-2(2) - Methyloccaceae bacterium NSP1-2, accession OYV21463.1; Crenothrix(1) - Crenothrix polyspora, accession WP_087142071.1; Crenothrix(2) - Crenothrix polyspora, accession WP_087147734.1; Methylocystis 49242 – Methylocystis sp. ATCC 49242, accession WP_036288389.1; Methylocystis SC2 - Methylocystis sp. SC2, accession WP_014891803.1; Methylocystis LW5 - Methylocystis sp. LW5, accession WP_026600015.1; Methylocystis bryophila - Methylocystis bryophila, accession WP_085772550.1; Methylocystis parvus - Methylocystis parvus OBBP, accession WP_020372486.1; Methylosinus LW4 - Methylosinus sp. LW4, accession WP_018266942.1; Methylosinus PW1 - Methylosinus sp. PW1, accession WP_084570776.1; Methylosinus R-45379 - Methylosinus sp. R-45379, accession WP_026600015.1; Methylosinus 3S-1 - Methylosinus sp. 3S-1, accession WP_003608458.1; Methylosinus LW3 - Methylosinus sp. LW3, accession WP_024878487.1; Pleomorphomonas koreensis - Pleomorphomonas koreensis DSM 23070, accession WP_026782666.1; Pleomorphomonas oryzae - Pleomorphomonas oryzae DSM 16300, accession WP_051228856.1; Methylocapsa - Methylocapsa palarum, accession WP_091682999.1; Methylohalobius 10Ki - Methylohalobius crimeensis 10Ki, accession WP_022947153.1. It is important to note that the Csp3 sequences from M. trichosporium OB3b and Methylosinus 3S-1, Methylosinus sp. LW4 and Methylosinus
sp. PW1, and *Methylocystis* sp. LW5 and *Methylosinus* sp. R-45379 are identical. The sequence of the Csp3 from *Methylosinus* sp. LW3 exhibits only a single difference to those of *Methylocystis* sp. LW5 and *Methylosinus* sp. R-45379. This very high sequence conservation is also true for MbnAs. The sequences from *M. trichosporium* OB3b and *Methylosinus* 3S-1 are identical, as are those from *Methylosinus* sp. PW1, *Methylocystis* sp. LW3 and *Methylosinus* sp. R-45379, with the MbnA from *Methylocystis* sp. LW5 having a single difference in its sequence.

References