

A new type of carboxysomal carbonic anhydrase in sulfur chemolithoautotrophs from alkaline environments

Jana Wieschollek,¹ Daniella Fuller,¹ Arin Gahramanova,¹ Terrence Millen,¹ Ashianna J. Mislav,¹ Ren R. Payne,¹ Daniel P. Walsh,¹ YuXuan Zhao,¹ Madilyn Carney,¹ Jaden Cross,¹ John Kashem,¹ Ruchi Korde,¹ Christine Lacy,¹ Noah Lyons,¹ Tori Mason,¹ Kayla Torres-Betancourt,¹ Tyler Trapnell,¹ Clare L. Dennison,¹ Dale Chaput,¹ Kathleen M. Scott¹

AUTHOR AFFILIATION See affiliation list on p. 19.

ABSTRACT Autotrophic bacteria are able to fix CO₂ in a great diversity of habitats, even though this dissolved gas is relatively scarce at neutral pH and above. As many of these bacteria rely on CO₂ fixation by ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) for biomass generation, they must compensate for the catalytical constraints of this enzyme with CO₂-concentrating mechanisms (CCMs). CCMs consist of CO₂ and HCO₃⁻ transporters and carboxysomes. Carboxysomes encapsulate RubisCO and carbonic anhydrase (CA) within a protein shell and are essential for the operation of a CCM in autotrophic *Bacteria* that use the Calvin-Benson-Basham cycle. Members of the genus *Thiomicrospira* lack genes homologous to those encoding previously described CA, and prior to this work, the mechanism of function for their carboxysomes was unclear. In this paper, we provide evidence that a member of the recently discovered iota family of carbonic anhydrase enzymes (ICA) plays a role in CO₂ fixation by carboxysomes from members of *Thiomicrospira* and potentially other *Bacteria*. Carboxysome enrichments from *Thiomicrospira pelophila* and *Thiomicrospira aerophila* were found to have CA activity and contain ICA, which is encoded in their carboxysome loci. When the gene encoding ICA was interrupted in *T. pelophila*, cells could no longer grow under low-CO₂ conditions, and CA activity was no longer detectable in their carboxysomes. When *T. pelophila* ICA was expressed in a strain of *Escherichia coli* lacking native CA activity, this strain recovered an ability to grow under low CO₂ conditions, and CA activity was present in crude cell extracts prepared from this strain.

IMPORTANCE Here, we provide evidence that iota carbonic anhydrase (ICA) plays a role in CO₂ fixation by some organisms with CO₂-concentrating mechanisms; this is the first time that ICA has been detected in carboxysomes. While ICA genes have been previously described in other members of bacteria, this is the first description of a physiological role for this type of carbonic anhydrase in this domain. Given its distribution in alkaliphilic autotrophic bacteria, ICA may provide an advantage to organisms growing at high pH values and could be helpful for engineering autotrophic organisms to synthesize compounds of industrial interest under alkaline conditions.

KEYWORDS carboxysome, alkaliphile, carbonic anhydrase, carbon dioxide concentrating mechanism, autotroph

A significant portion of global CO₂ fixation is catalyzed within carboxysomes (1). Carboxysomes are proteinaceous microcompartments that facilitate CO₂ fixation by many autotrophic bacteria that use the Calvin-Benson Bassham cycle [CBB (2)]. Carbon fixation by the CBB cycle is catalyzed by RubisCO (ribulose 1,5-bisphosphate carboxylase/oxygenase), which has a low affinity for CO₂, and can also use O₂ as a substrate for a wasteful side reaction (3). In order to compensate for the catalytic constraints of

Editor Jennifer F. Biddle, University of Delaware, Lewes, Delaware, USA

Address correspondence to Kathleen M. Scott, kmscott@usf.edu.

The authors declare no conflict of interest.

See the funding table on p. 19.

Received 30 May 2024

Accepted 1 August 2024

Published 23 August 2024

Copyright © 2024 American Society for Microbiology. All Rights Reserved.

RubisCO, many autotrophic bacteria have a CO₂-concentrating mechanism (CCM), which consists of two components: dissolved inorganic carbon (DIC; CO₂, HCO₃[−], and CO₃^{2−}) transporters, which transport DIC across the cell membrane to generate elevated levels of HCO₃[−] within the cytoplasm, and carboxysomes (4, 5). Carboxysomes contain two key enzymes: (i) carbonic anhydrase (CA), which interconverts HCO₃[−] and CO₂ and (ii) RubisCO, which uses CO₂ as its substrate to carboxylate ribulose 1,5-bisphosphate (6–9). When cytoplasmic HCO₃[−] enters carboxysomes, CA converts it to CO₂, and it is fixed by RubisCO (10).

Two types of carboxysomes are known: α-carboxysomes, which are present in autotrophic *Proteobacteria* and some *Cyanobacteria*, including marine globally abundant *Prochlorococcus* and *Synechococcus* spp. (1), and β-carboxysomes, which are found exclusively in *Cyanobacteria* (11–13). While the two types have the same function in CCMs, they arose independently and can be distinguished by the forms of RubisCO and CA they encapsulate (9, 12). α-Carboxysomes contain form IA RubisCO and a β-class CA [CsoSCA (8)], while β-carboxysomes contain form IB RubisCO and either β-class or γ-class CA or both (14–17). Form IA and form IB RubisCO share a common ancestry, while the different forms of carboxysomal CA (β and γ) are evolutionarily independent (18). Notably, these are not the only classes of CA that exist. CA consists of a superfamily of diverse enzymes and is present in all domains of life. Currently, there are eight (α, β, γ, δ, ζ, η, θ, and ι) classes of CA known, and at least six of these classes are evolutionary independent forms [reviewed in reference (5)]. All classes facilitate the reversible hydration of CO₂ to HCO₃[−], and most require metals for activity [commonly Zn⁺² (18)].

Given the uniformity of CA presence in biochemically characterized α- and β-carboxysomes, and the critical role CA plays in carboxysome function, it was surprising to find that sulfur-oxidizing autotrophic members of genus *Thiomicrospira*, from phylum *Gammaproteobacteria*, had carboxysome loci lacking genes encoding the CA typically associated with α-carboxysomes [CsoSCA; Fig. 1 (19, 20)]. In its place, they have a smaller gene, *csoSX*, which does not encode residues necessary for CsoSCA activity. These differences are interesting since other evidence suggests the presence of a typical CCM in members of *Thiomicrospira*. *Thiomicrospira pelophila* and *Thiomicrospira thyasirae*, like many other organisms with CCMs (19, 21), have genes encoding both carboxysomal RubisCO (form I; *cbbL* and *cbbS*), as well as a presumably non-carboxysomal form II RubisCO (*cbbM*). Other organisms with multiple RubisCO-encoding genes upregulate *cbbM* when CO₂ is abundant and upregulate carboxysomal *cbbL* and *cbbS* when CO₂ is scarce as part of their CCM (22, 23). Carboxysomes have previously been noted in electron micrographs of members of *Thiomicrospira* (24–26), and it was puzzling that these carboxysomes could function in an autotrophic cell in the absence of CA. Recently, an additional form of CA was discovered in phytoplankton [iota carbonic anhydrase (ιCA) (27)], and a homolog to ιCA is present in carboxysome loci in all members of genus *Thiomicrospira* (Fig. 1). So far, ιCA has only been described and biochemically characterized in a handful of organisms [*Thalassiosira pseudonana*, *Burkholderia territorii*, *Bigeloviella natans*, and *Anabaena* sp. PCC7120 (27–30)], and homologs are common in organisms from all three domains. In *Bacteria*, two ιCA have been biochemically characterized (28, 29), but their roles in the physiologies of these organisms have not been studied. Given the presence of genes encoding ιCA homologs in carboxysome loci from multiple organisms, our hypothesis is that ιCA functions as the carboxysomal CA in these organisms, the first indication that ιCA may play a role in these microcompartments.

RESULTS

Purified carboxysomes from *T. pelophila* and *Thiomicrospira aerophila*

Carboxysomes were successfully enriched from *T. pelophila* and *T. aerophila* grown under low-DIC conditions (Table 1). Suspensions of carboxysomes were dominated by carboxysome proteins, as determined by LC-MS/MS (Table S1). A small number of major bands are visible in Coomassie blue-stained SDS-PAGE gels, and these bands

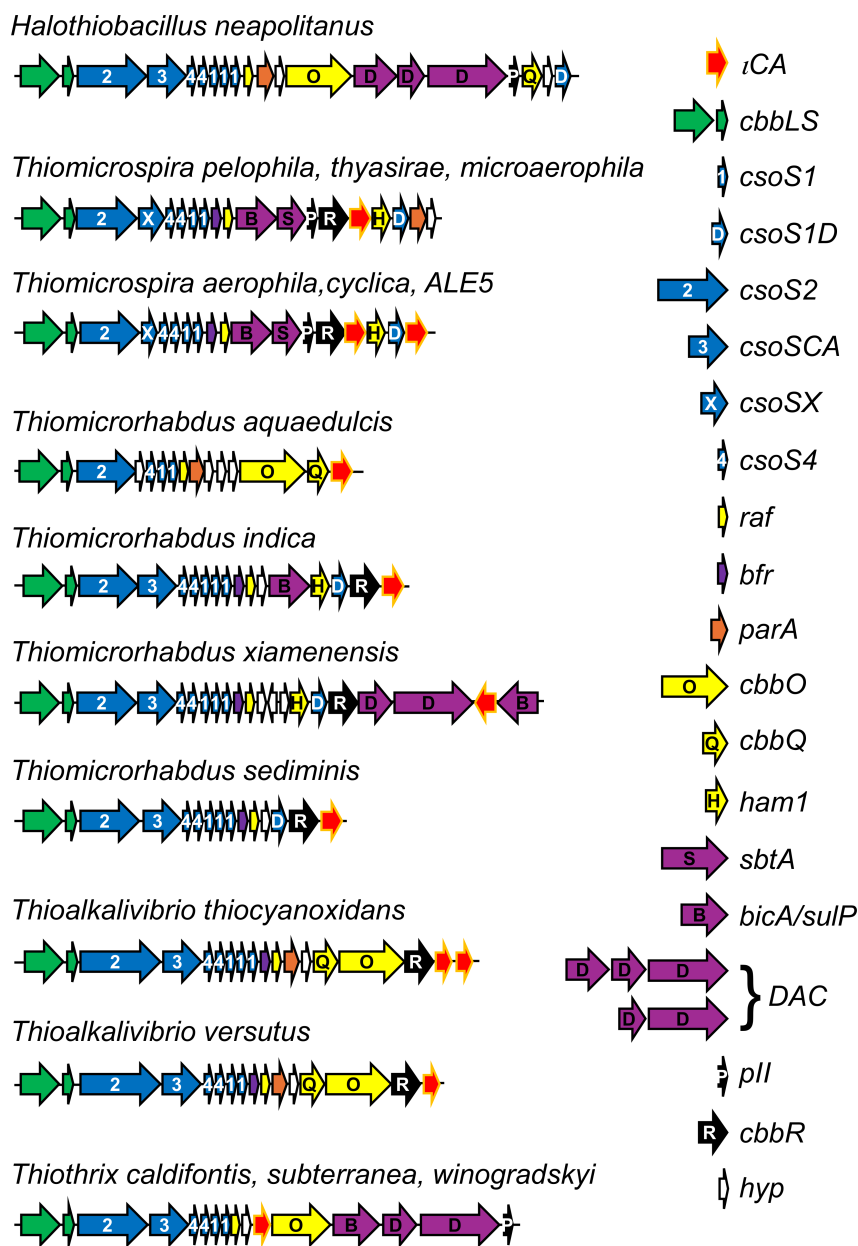


FIG 1 Carboxysome loci of members of *Thiomicrospira* and others that include *iCA* homologs. The carboxysome locus from *Halothiobacillus neapolitanus*, which is a model organism for carboxysomes in *Proteobacteria*, is included for comparison. For genera *Thioalkalivibrio* and *Thiothrix*, many strains have carboxysome loci with *iCA* homologs nearby, and representative strains are included in this figure. Carboxysome loci with *iCA* homologs were found via cassette cluster search of the Integrated Microbial Genomes and Microbiomes database (<https://img.jgi.doe.gov/>), to find members of Pfams 08332 (*iCA*), 12288 (*CsoS2*), and 00101 (*CbbS*, RubisCO small subunit) within 20 kbp of each other on a genome sequence. Abbreviations: *bfr*, bacterioferritin; *cbbLS*, large and small subunits of RubisCO; *cbbO* and *cbbQ*, RubisCO activases (31); *cbbR*, LysR-type regulatory protein (32); *csoS1*, *S1D*, and *S4*, carboxysome shell proteins (33–35); *csoS2*, carboxysome assembly protein (36); *csoSCA*, carboxysomal carbonic anhydrase (8); *csoSX*, homologous to the amino terminus of *csoSCA*, lacking CA active-site residues encoded in *csoSCA*; *ham1*, Ham1 family protein (Pfam01725); *iCA*, iota carbonic anhydrase homologs; *pII*, signal transduction protein (37); *parA*, plasmid partitioning protein homolog (Pfam01656); *raf*, RubisCO assembly factor (38); *sbtA*, *bicA/sulP*, and *DAC*, DIC transporters (19).

were excised and identified as carboxysome proteins via LC-MS/MS (Fig. 2A). For *T. aerophila* carboxysome preparations, porin was present, indicating some contamination from the outer membrane. For both species, α CA was detected at low abundance in some carboxysome preparations ($\leq 0.025\%$ of signal intensity; Table 1) and was not enriched in carboxysome preparations. Intact and broken carboxysomes were visible via transmission electron microscopy of negatively stained carboxysome preparations (Fig. 2B).

CA activity in *T. pelophila* and *T. aerophila* carboxysomes

CA activity was present in purified carboxysomes that had been freeze-thawed to release the enzymes they contain [Fig. 3; Table 2 (39)]. For carboxysomes from either *T. pelophila* or *T. aerophila*, boiled samples had significantly lower rates of CO_2 hydration due to chemical (non-enzymatic) CO_2 hydration remaining after heat treatment ($P < 0.005$; Table 2). CA activities, calculated as the difference between boiled and unboiled rates, were similar for carboxysomes from both organisms (Table 2). *T. pelophila* carboxysome-associated CA was inhibited by ethoxzolamide (EZA) dissolved in dimethyl sulfoxide (DMSO) relative to the solvent control (EZA vs DMSO solvent control, $P < 0.015$; the addition of DMSO itself did not inhibit CA activity, $P > 0.96$). CA activity from *T. aerophila* carboxysomes was lower in the presence of EZA, though not to statistical significance ($P < 0.1$,

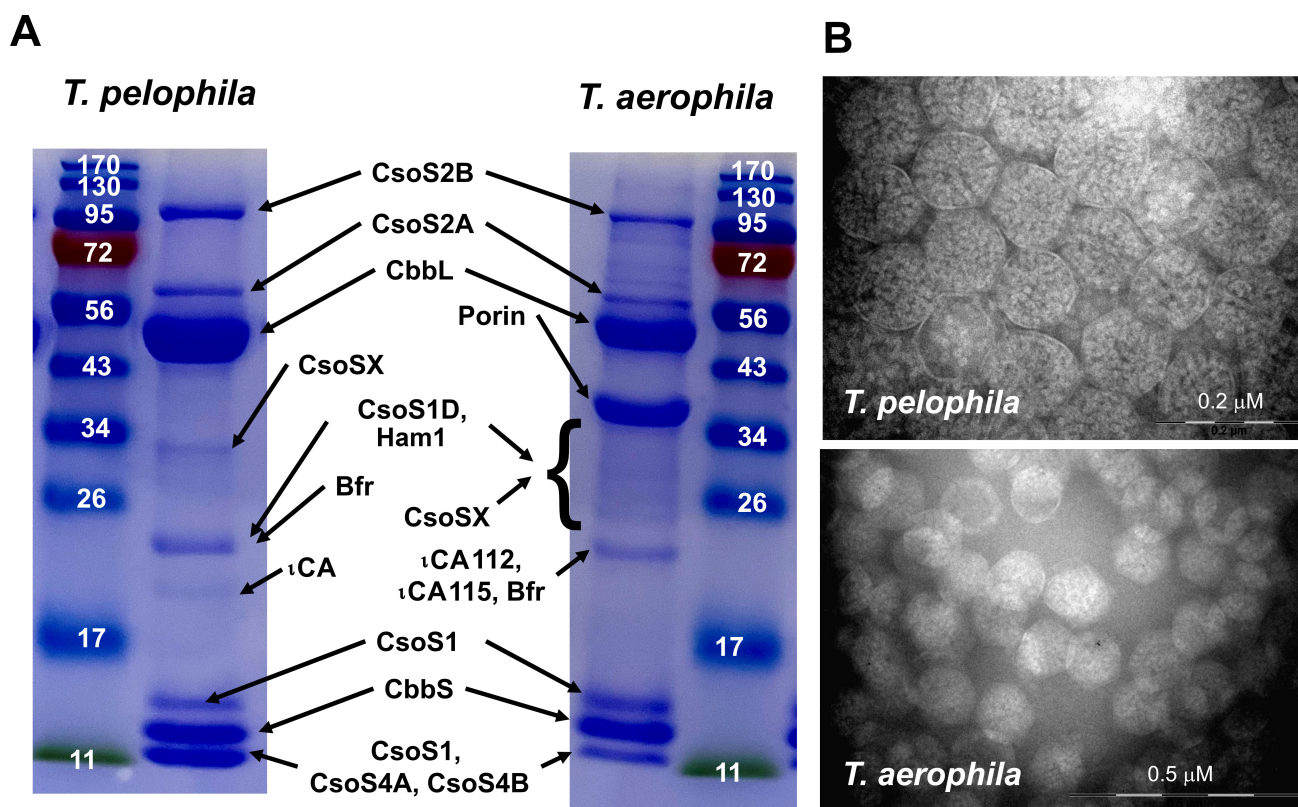


FIG 2 Carboxysome preparations from wild-type *T. pelophila* and *T. aerophila*. (A) SDS-PAGE analysis of carboxysome preparations from *T. pelophila* and *T. aerophila* (100 μg protein was loaded for each; protein ladder is labeled in kDa). Proteins from the major bands were identified via LC-MS/MS, including CsoS2A and B (Fig. S1). Two bands are apparent for CbbS; the one labeled CbbS* may be post-translationally modified (please see Fig. S2 and S3, and text for further information on bands A–C). Abbreviations: Bfr, bacterioferritin; CbbL and CbbS, large and small subunits of RubisCO; CsoS1, S1D, and S4, carboxysome shell proteins (33–35); CsoS2, carboxysome assembly protein (36); CsoSX, homologous to the amino terminus of CsoSCA, lacking CA active-site residues found in CsoSCA; Ham1, Ham1 family protein (Pfam01725); α CA, α -carbonic anhydrase homologs. (B) Transmission electron micrographs of carboxysomes negatively stained with ammonium molybdate. Carboxysomes from *T. pelophila* were magnified to 140,000 $\times$, and carboxysomes from *T. aerophila* were magnified to 100,000 $\times$.

TABLE 1 Relative abundance of carboxysome-associated proteins in carboxysome preparations

Proteins	% of total signal intensity ^a			
	<i>T. pelophila</i> carboxysomes (n = 2)	<i>T. pelophila</i> cells (n = 3)	<i>T. aerophila</i> carboxysomes (n = 2)	<i>T. aerophila</i> cells (n = 3)
Cso + CbbL + CbbS	82.9, 37.6	5.2 ± 0.2	41.5, 42.4	6.6 ± 0.3
CbbL	32.8, 17.2	1.90 ± 0.06	17.5, 15.8	2.43 ± 0.13
CbbS	11.2, 3.6	0.44 ± 0.07	6.77, 7.01	0.76 ± 0.22
CsoS2	10.7, 8.6	1.19 ± 0.08	7.3, 10.3	1.49 ± 0.04
CsoSX	0.99, 0.20	0.020 ± 0.001	0.10, 0.13	0.0054 ± 0.0014
CsoS4A	1.23, 0.17	0.027 ± 0.005	0.17, 0.19	0.0024 ± 0.0008
CsoS4B	0.17, 0.03	0.003 ± 0.002	0.04, 0.10	0.0040 ± 0.0023
CsoS1 (A-C)	23.6, 7.6	1.63 ± 0.08	9.31, 8.40	0.086 ± 0.079
CsoS1D	2.18, 0.29	0.044 ± 0.003	0.35, 0.45	0.040 ± 0.003
Bfr ^b	6.91, 1.58	0.021 ± 0.03	1.31, 1.07	0.11 ± 0.01
Ham1 ^c	0.59, 0.10	0.016 ± 0.002	0.15, 0.18	0.035 ± 0.013
iCA	3.53, 0.40	0.063 ± 0.007	–	–
iCA 112 ^d	–	–	0.07, 0.05	0.0049 ± 0.0007
iCA 115 ^e	–	–	0.95, 1.23	0.14 ± 0.01
αCA ^f	0.025, ND ^g	0.027 ± 0.002	ND, 0.0049	0.0075 ± 0.0002

^a% of signal intensity for all proteins detected in carboxysome samples or cells by LC-MS/MS.

^bBacterioferritin; member of Pfam00210.

^cHam1 family protein, XTP/dITP diphosphohydrolase; member of Pfam01725.

^diCA 112 is encoded in *T. aerophila* by IMG gene object identifier 2507134112.

^eiCA 115 is encoded in *T. aerophila* by IMG gene object identifier 2507134115.

^fαCA is encoded by IMG gene object identifier 2568511177 in *T. pelophila* and 2507133695 in *T. aerophila*.

^gND, none detected.

EZA vs DMSO, $P > 0.7$, DMSO vs no additions). CA activity from *T. aerophila* carboxysomes was sensitive to dithiothreitol ($P < 0.005$), while CA from *T. pelophila* was not ($P < 0.56$).

Characterization of CO₂-sensitive *T. pelophila* mutants

From the library of 10,000 mutant strains of *T. pelophila*, 19 were unable to grow under low-CO₂ conditions (Fig. S4). Because the initial screen of the library was a growth/no growth assay under low-CO₂ conditions, it is possible that other strains in the library, beyond the 19 studied further here, were capable of lower rates of growth under low-CO₂ conditions. In these 19 strains, Tn5-RL27 was inserted within the carboxysome operon (Fig. 4A). Some of these insertions fell within the genes encoding carboxysomal RubisCO (*cbbL*; Fig. 4A). Presumably, these mutant strains were still capable of growth under high-CO₂ conditions via expression of *cbbM*, which encodes non-carboxysomal form II RubisCO encoded elsewhere in the genome. Carboxysomes were not apparent in mutants whose Tn5-RL27 insertion sites fell within genes encoding shell-associated proteins (*csoS2* and *csoSX*; Fig. 4C and D). However, mutants with an interruption within the gene encoding the iCA homolog were able to generate intact carboxysomes (Fig. 4E and F).

Growth and CA activities of *Escherichia coli* constructs expressing genes from *T. pelophila*

E. coli Lemo21(DE3) $\Delta yadF \Delta cynT$ successfully expressed CsoS2, CsoSX, and iCA from *T. pelophila*, based on the presence of these proteins in cell pellets (Table S2). Cells expressing CsoS2 or CsoSX were unable to grow under low-CO₂ conditions despite growing robustly under a 5% CO₂ headspace (Fig. 5). When expressing the iCA homolog from *T. pelophila*, *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ was able to grow under ambient air (0.04% CO₂; Fig. 5). This construct had measurable CA activity 6 hours after inducer was added (6 hours: $0.53 \pm 0.02 \mu\text{moles CO}_2 \text{ s}^{-1} \text{ mg protein}^{-1}$, $P < 0.002$; 8 hours: $0.36 \pm 0.01 \mu\text{moles CO}_2 \text{ s}^{-1} \text{ mg protein}^{-1}$, $P < 0.05$; 24 hours: $0.49 \pm 0.03 \mu\text{moles CO}_2 \text{ s}^{-1} \text{ mg protein}^{-1}$,

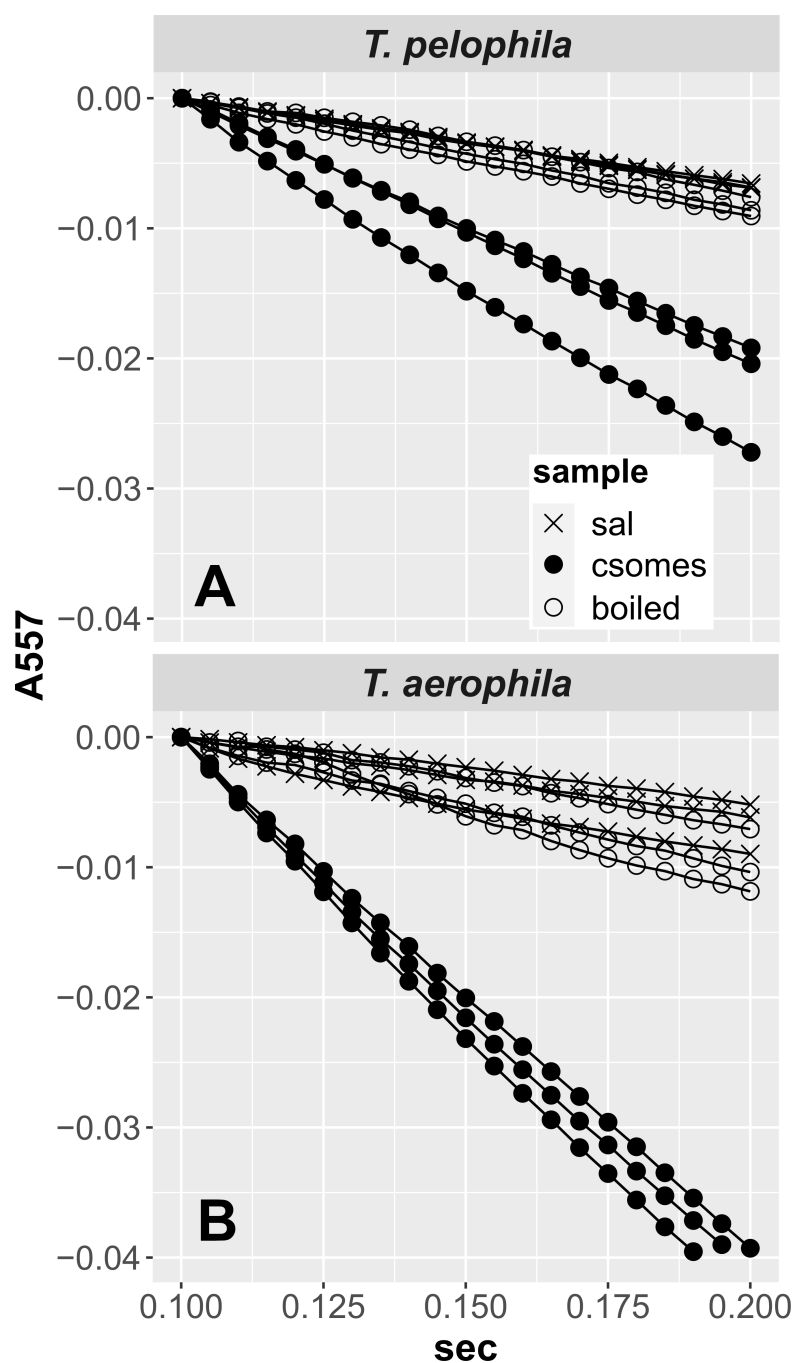


FIG 3 CO₂ hydration by carboxysomes purified from *T. pelophila* (A) and *T. aerophila* (B). CO₂ hydration was measured indirectly via the response of phenol red to assay acidification by protons released by carbonic acid formed from CO₂. CO₂ hydration was measured in three samples of saline (sal), carboxysomes (csomes), and boiled carboxysomes (boiled). Absorbances were calculated by subtracting the initial absorbance value (A_0) from each timepoint absorbance value (A_t ; $A_{557} = A_t - A_0$) to compensate for turbidity due to the presence of carboxysomes.

$P < 0.05$). CA activity was not detected in *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ in the absence of genes from *T. pelophila* ($P > 0.05$ at all timepoints).

TABLE 2 Activity and inhibition of carboxysome-associated carbonic anhydrase

Organism	mg protein mL ⁻¹	Carbonic anhydrase activity (μmoles CO ₂ s ⁻¹ mg protein ⁻¹) ±SD (n = 3)			
		No inhibitors	10 mM DTT ^a	1% (vol/vol) DMSO ^b	1 mM EZA ^c , 1% (vol/vol) DMSO
<i>T. pelophila</i>	0.7	3.1 ± 0.9	3.9 ± 0.3	3.5 ± 1.0	1.5 ± 0.4
<i>T. aerophila</i>	2.3	2.1 ± 0.3	0 ± 0.2	2.0 ± 0.2	1.6 ± 0.3

^aDTT, dithiothreitol.

^bDMSO, dimethyl sulfoxide.

^cEZA, ethoxzolamide; 6-ethoxy-2-benzothiazolesulfonamide.

Characterization of carboxysomes purified from *T. pelophila* with a mutation in the gene encoding iCA

Carboxysomes from mutant *T. pelophila* strain iCA-135 were successfully purified. When subject to SDS-PAGE analysis, a small number of major bands were visible in Coomassie blue-stained gels (Fig. 6A), and LC-MS/MS analysis confirmed that carboxysome proteins comprised 92% of all proteins in the sample (Table S1). Intact and broken carboxysomes were visible in transmission electron micrographs (Fig. 6B). Relative amounts of many carboxysome proteins were similar in mutant and wild-type carboxysomes. However, amounts of iCA and CsoS1D were lower in two independent preparations of mutant carboxysomes (Fig. 6C).

The iCA present in mutant carboxysomes consists of the amino terminus of the protein (Fig. 6D), which reflects the insertion site of the transposon. The diminishment in CsoS1D abundance likely reflects the polar effects of transposon insertion in the gene encoding iCA, which precedes the gene encoding CsoS1D (Fig. 1). Additionally, one of the lower molecular weight bands from wild-type preparations (Fig. 2; labeled A–C) is absent in mutant carboxysomes (Fig. 6A; labeled B and C). The identities of the proteins in these bands were confirmed by LC-MS/MS analysis of bands excised from SDS-PAGE gels. The composition of bands B and C from mutant carboxysomes resembles the compositions of their counterparts from wild-type cells (Fig. S2), suggesting the loss of the proteins comprising band A from wild-type carboxysomes in mutant carboxysomes. Given that band A in wild-type carboxysome preparations is similar in composition to band B and consists primarily of CbbS (Fig. S2), it is possible that CbbS (or CsoS1.1 and CsoS1.2, which are also present in band A) is post-translationally modified in wild-type cells. The abundance of peptide fragments generated during LC-MS/MS from CbbS, CsoS1.1, and CsoS1.2 was compared (Fig. S3), figuring that post-translational modification of the intact protein at a particular amino acid residue could change the molecular weight and/or charge of the fragment, resulting loss of detection. Indeed, protein fragments corresponding to CbbS (peptide corresponding to residues 70–83), CsoS1.1 (peptides corresponding to residues 1–15 and 39–51), and CsoS1.2 (same as CsoS1.1) are absent in wild-type band A, which is consistent with potential post-translational modification of these proteins (Fig. S3). It is possible that the gene necessary for post-translational modification of these proteins is encoded downstream from iCA, and its expression was diminished similarly to CsoS1D.

CA activity was not detected in the mutant carboxysomes (Fig. 6E). There was no significant difference in CA activity between the carboxysome suspension and the boiled control ($P > 0.12$). The concentration of protein in the assay was 0.4 mg mL⁻¹, lower than the assay of wild-type carboxysomes (0.7 mg mL⁻¹). However, carboxysome proteins were more enriched in preparations of mutant carboxysomes (92%) than for the preparation from wild-type *T. pelophila* used for this assay (38%), suggesting that the negative result is not a dilution effect. These carboxysomes had RubisCO activities (0.3 μmol CO₂/min mg protein) similar to wild-type carboxysomes (two independent preparations of wild-type carboxysomes: 0.8 and 0.4 μmol min⁻¹ mg⁻¹ protein, respectively), suggesting that the absence of CA activity was not due to carboxysome breakage and loss of contents (e.g., RubisCO, CA) during purification.

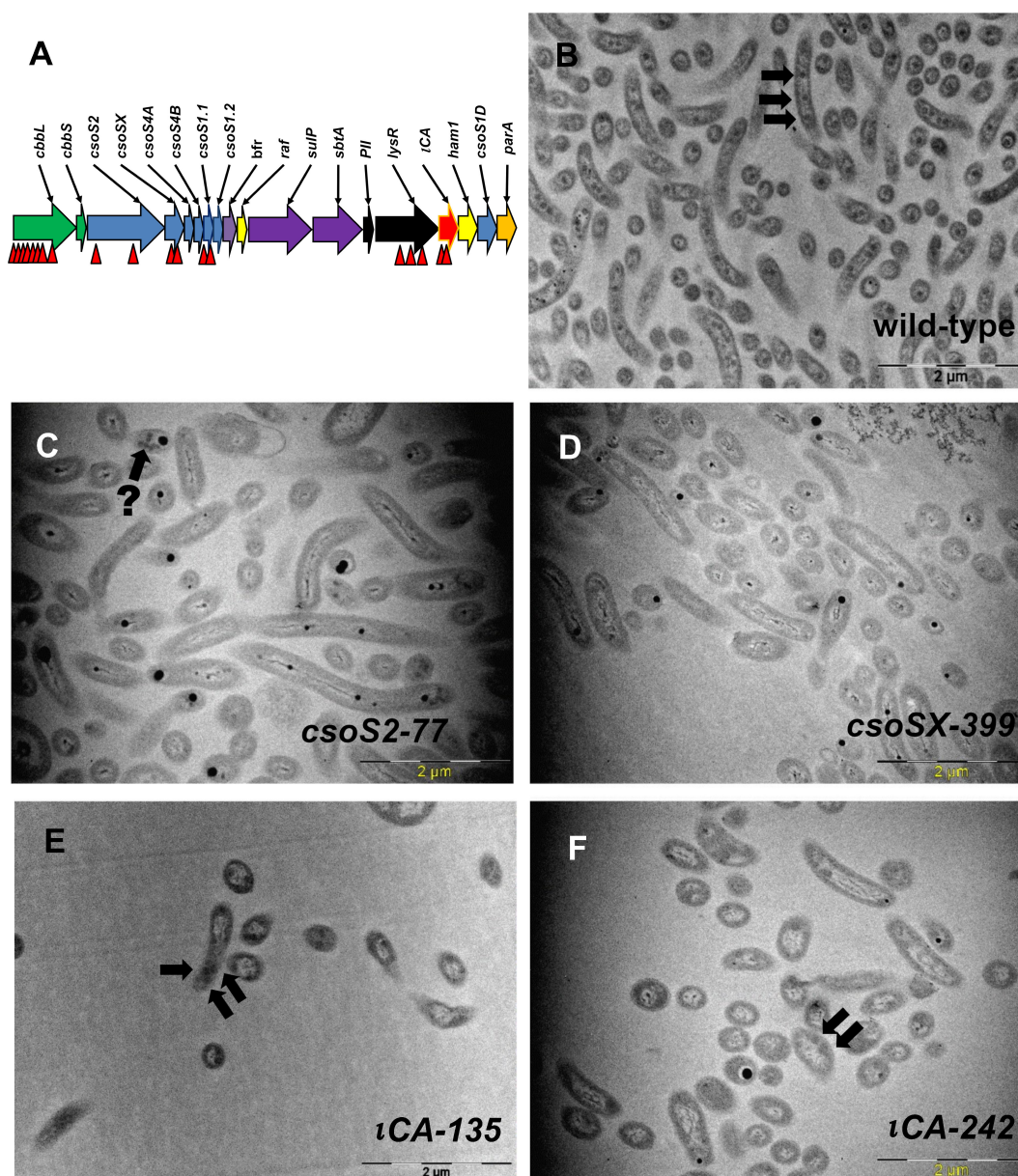


FIG 4 CO₂-sensitive *T. pelophila* strains generated by transposon-mediated random mutagenesis. (A) Transposon insertion sites (red arrows) present in the CO₂-sensitive strains (one per strain); exact locations of insertion sites are indicated in Fig. S4. Transmission electron micrographs (14,000× magnification) of wild-type *T. pelophila* (B) and CO₂-sensitive mutants (C–F) with arrows indicating carboxysomes, and numbers after gene names indicating the location of the transposon insertion site in nucleotides from the first nucleotide of the start codon of the targeted gene. In (C), a cell has inclusions marked with (?), which appear to be slightly smaller than wild-type carboxysomes. Abbreviations: bfr, bacterioferritin; cbbL and cbbS, large and small subunits of RubisCO; cbbR, LysR-type regulatory protein (32); csoS1, S1D, and S4, carboxysome shell proteins (33–35); csoS2, carboxysome assembly protein (36); cbbO and cbbQ, RubisCO activases (31); csoSX, homologous to the amino terminus of csoSCA, lacking CA active site residues encoded in csoSCA; ham1, Ham1 family protein (Pfam01725); iCA, iota carbonic anhydrase homologs; pII, signal transduction protein (37); parA, plasmid partitioning protein homolog (Pfam01656); raf, RubisCO assembly factor (38); sbtA, bicA/sulP, and DAC, DIC transporters (19).

Phylogenetic analysis of iCA homologs from genus *Thiomicrospira*

Genes encoding iCA from *Thiomicrospira* spp. fall among those from other members of *Gammaproteobacteria*, divergent from those previously described and biochemically characterized (Fig. 7). For those *Thiomicrospira* species with two iCA homologs, these genes fall into two separate clusters.

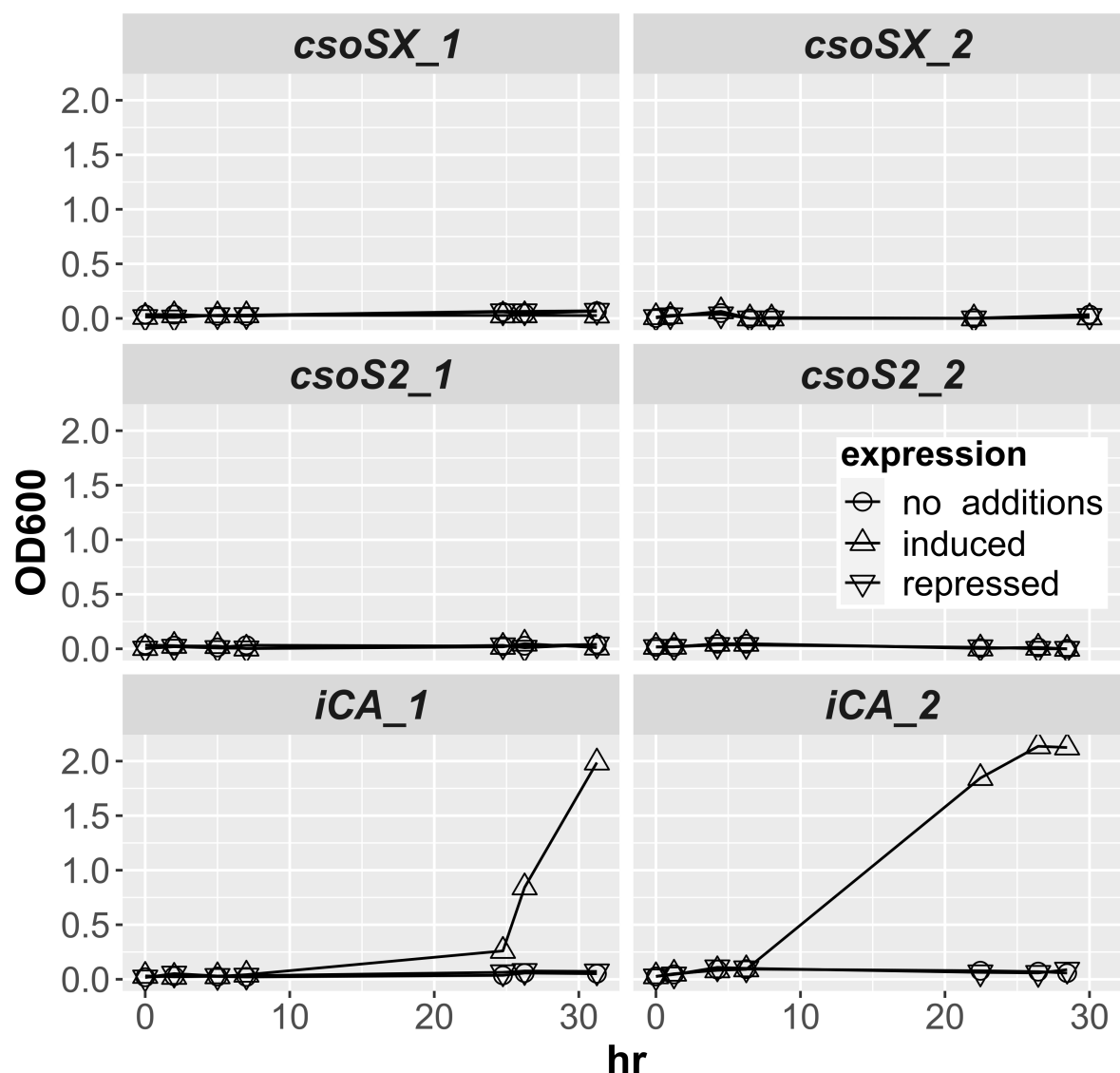


FIG 5 Growth curves of *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ (double carbonic anhydrase knockout) expressing genes required for low CO_2 growth by *T. pelophila*. Duplicate cultures (1, 2) expressing genes controlled by the P_{BAD} promoter were grown. Top: *csoSX* (homologous to the amino terminus from *csoSCA*); middle: *csoS2* (shell assembly protein); bottom: homolog of *iCA*. Induced cells were cultivated in a medium supplemented with 6 mM arabinose, while the growth medium for repressed cells was supplemented with 11 mM glucose.

DISCUSSION

iCA is likely to function as the carboxysomal CA in the genus *Thiomicrospira*

Multiple lines of evidence presented here are consistent with the displacement of CsoSCA, formerly the only CA associated with α -carboxysomes, with *iCA* in carboxysomes from members of *Thiomicrospira*. Genes encoding *iCA* homologs are present in *Thiomicrospira* carboxysome loci, which lack genes encoding CsoSCA for all members of this genus sequenced thus far (Fig. 1). Though CsoSCA is absent, *Thiomicrospira* carboxysomes have CA activity (Table 2; Fig. 3), *iCA* proteins are present in preparations of purified carboxysomes (Table 1; Table S1), interruption of the gene encoding *iCA* results in loss of CA activity in carboxysomes (Fig. 6), and heterologous expression of *iCA* genes from *T. pelophila* in *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ (double CA knockout) rescues the strain's ability to grow under low CO_2 conditions (Fig. 5) and results in measurable CA

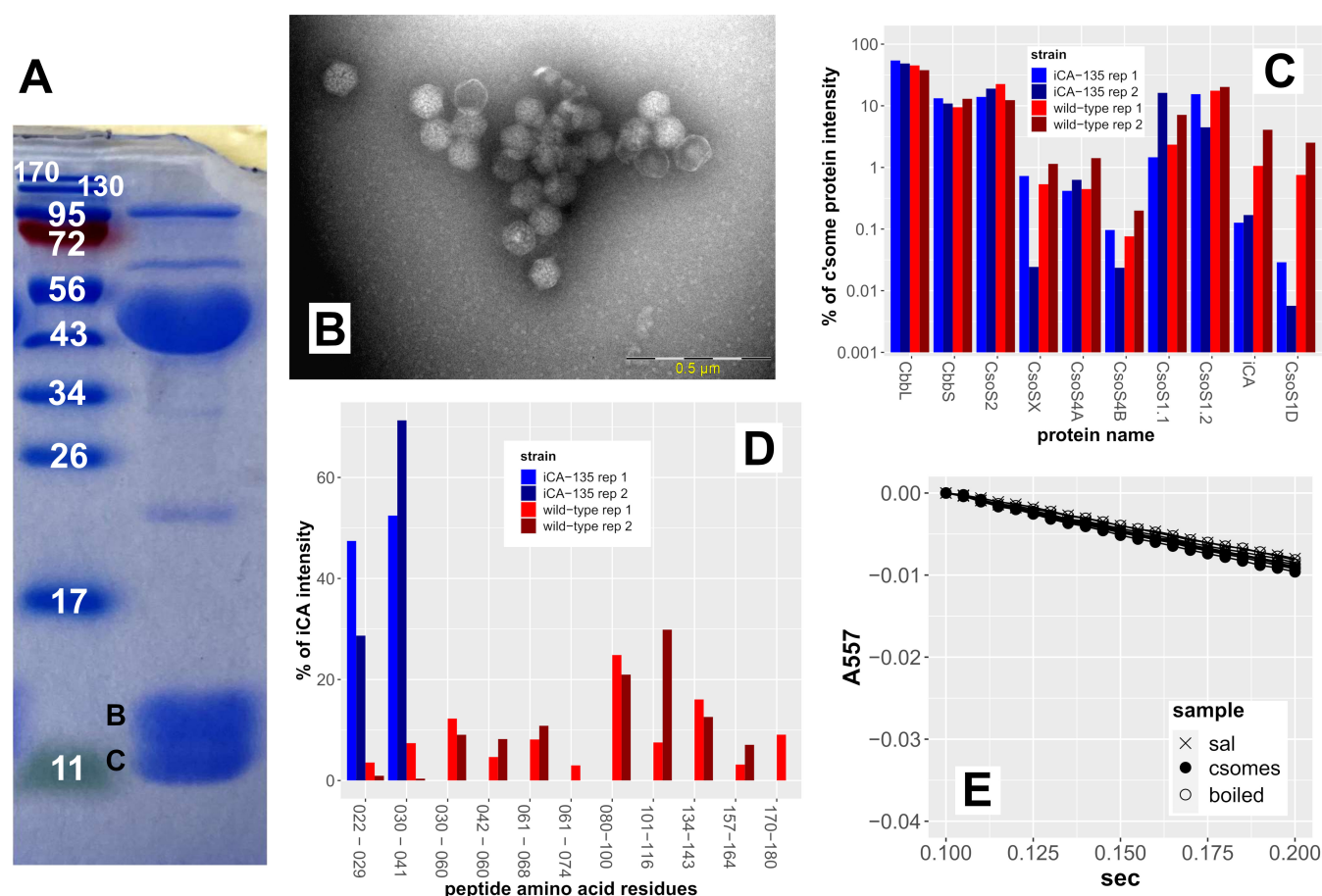


FIG 6 Carboxysomes from *T. pelophila* iCA-135. (A) SDS-PAGE of a carboxysome preparation from *T. pelophila* iCA-135 (100 μ g protein; protein ladder labeled in kDa). (B) Transmission electron micrographs of carboxysomes from *T. pelophila* iCA-135 negatively stained with ammonium molybdate. Carboxysomes were magnified 56,000 $\times$. (C) Relative amounts of carboxysome proteins in two independent replicate carboxysome preparations from wild-type *T. pelophila* (wild-type rep 1 and wild-type rep 2) and *T. pelophila* iCA-135 (iCA-135 rep 1 and iCA-135 rep 2). % of csome proteins is calculated from the intensity of LC-MS/MS signal from a particular protein (e.g., CbbL) divided by the total intensity of LC-MS/MS signal from all carboxysome-associated proteins (CbbL + CbbS + CsoSX + CsoS4A + CsoS4B + CsoS1.1 + CsoS1.2 + iCA + CsoS1D). Abbreviations: CbbL, large subunit of RubisCO; CsoS1.1, CsoS1.2, CsoS1D, CsoS4A, and CsoS4B, carboxysomal shell proteins; CsoS2, carboxysomal shell assembly protein; CsoS4a CsoSX, homologous to the amino terminus of csoSCA, lacking CA active-site residues encoded in csoSCA; iCA, homolog of iCA. (D) Signal intensities from iCA-derived peptides from two independent replicate carboxysome preparations from wild-type *T. pelophila* (wild-type rep 1 and wild-type rep 2) and *T. pelophila* iCA-135 (iCA-135 rep 1 and iCA-135 rep 2). Signal intensities from peptides resulting from iCA protein fragmentation for LC-MS/MS analysis were summed and used as the denominator to calculate the percent of signal from individual iCA peptides. Numbers on the x-axis correspond to the location of the peptide in the complete iCA sequence. (E) CO₂ hydration by carboxysomes isolated from *T. pelophila* iCA-135. CO₂ hydration was measured indirectly via the response of phenol red to assay acidification by protons released by carbonic acid formed from CO₂. CO₂ hydration was measured in three samples of saline (sal), carboxysomes (csomes), and boiled carboxysomes (boiled). Absorbances were calculated by subtracting the initial absorbance value (A_0) from each timepoint absorbance value (A_t ; $A_{557} = A_t - A_0$) to compensate for turbidity due to the presence of carboxysomes.

activity. To our knowledge, this is the first study to show the replacement of CsoSCA in α -carboxysomes, as well as a physiological function for iCA in *Bacteria*.

iCA enzymes from members of the genus *Thiomicrospira* share many similarities to those previously described in other *Bacteria*. Residues corresponding to those present in the active site of iCA from *Nostoc* sp. PCC7120 [T106, Y127, H197, and S199; HHSS motif at the carboxy terminus (29)] are present in iCA from *Thiomicrospiras* (Fig. S5). Similarities extend to predicted tertiary structure. Structures predicted via AlphaFold2 (43) are similar to the structure empirically determined for iCA from *Nostoc* sp. PCC7120 [Fig. 8 (29)].

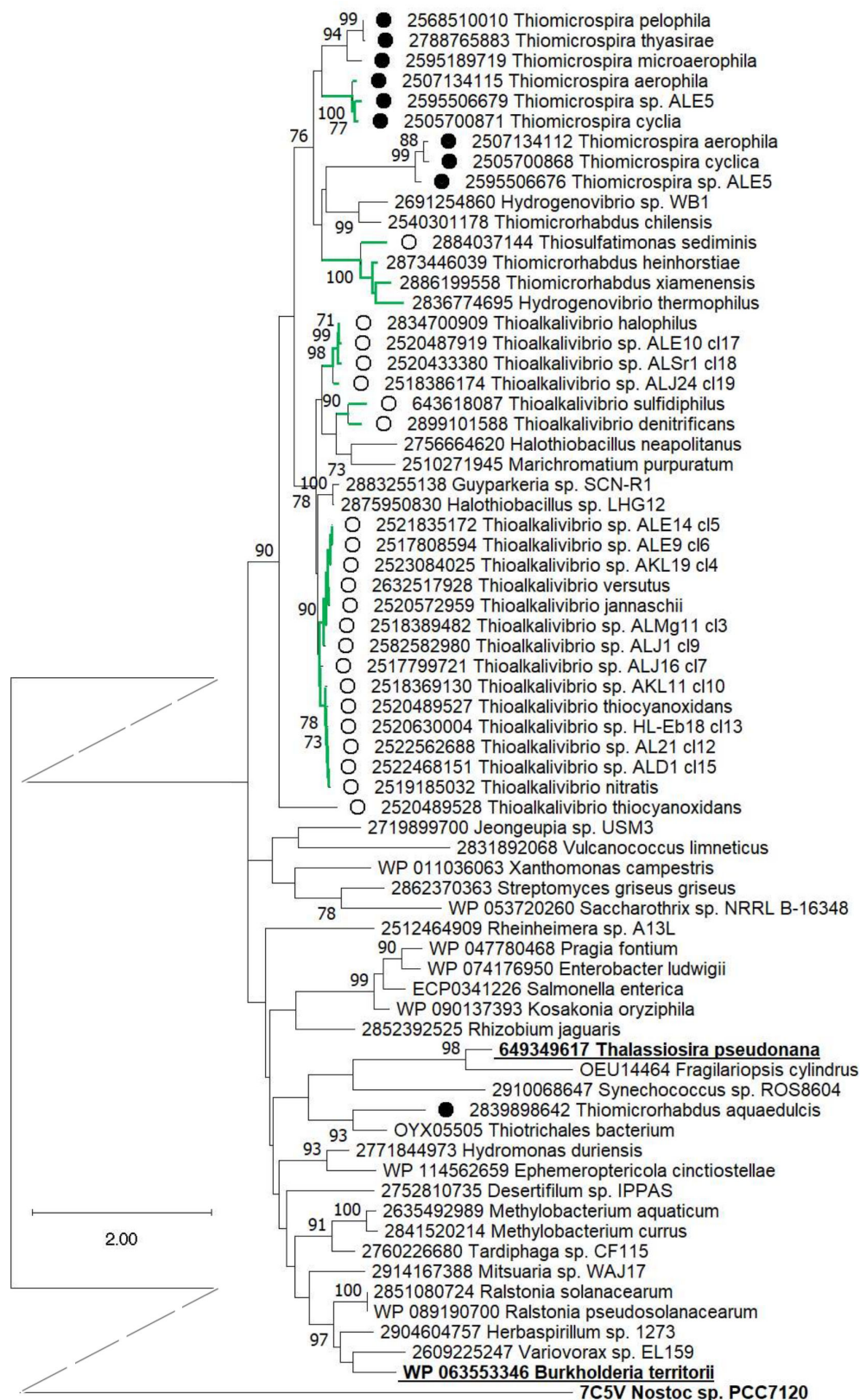


FIG 7 Maximum likelihood analysis of iCA amino acid sequences. Taxon names of biochemically characterized iCA are underlined (28, 29, 40). ●, carboxysome locus-associated, with *csdX* in the carboxysome locus; ○, carboxysome locus-associated, with *csdA* in the carboxysome locus; green-thickened branches, amino acid sequences with multiple cysteine residues (Continued on next page)

FIG 7 (Continued)

near the amino terminus. Numbers preceding the taxon names are gene object ID numbers from IMG (<https://img.jgi.doe.gov/> (41)) and accession numbers from the nonredundant protein sequences database (<https://www.ncbi.nlm.nih.gov/>). Sixty-nine sequences with 291 positions were used to generate the tree using the Le_Gascuel_2008 model (42) and a gamma distribution with five categories. The scale indicates the numbers of substitutions per site. Bootstrap values >70 are shown and are based on 100 resamplings of the alignment.

CsoSX is not a carboxysomal CA

T. pelophila CsoSX presence does not result in the growth of *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ under low CO₂ conditions (Fig. 5), consistent with an absence of CA activity. CsoSX sequences from *Thiomicrospiras* (250–332 aa) are substantially shorter than CsoSCA sequences (471–531 aa) and lack all of the residues necessary for CA activity by CsoSCA (Fig. S6; residues corresponding to C173, H242, and C253 in *H. neapolitanus*) (18, 46). Predicted structures for CsoSX proteins (Fig. 9A) are low-confidence and do not resemble the empirically determined structure for CsoSCA from *H. neapolitanus* [PDB 2FGY (46); root mean square deviation of distances is high, >6 Å, and template modeling scores are low, ~0.2].

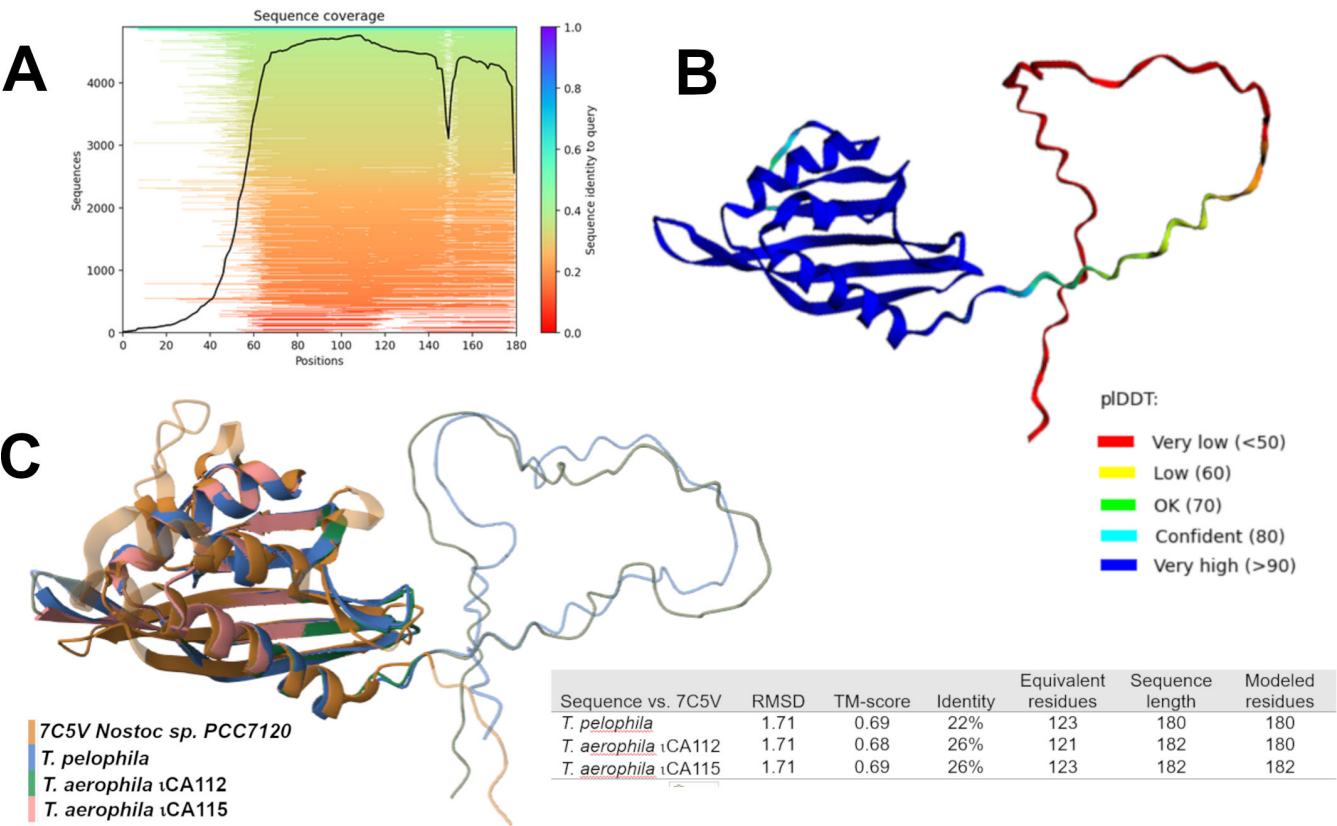


FIG 8 Structures predicted for iCA homologs from members of genus *Thiomicrospira*. (A and B) Predicted structure of iCA from *T. pelophila* via AlphaFold2 (43), implemented via Colabfold (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) using default parameters, without comparison to Protein Data Bank (PDB) structures. (A) Comparison of iCA sequence from *T. pelophila* to homologs used to predict its structure. (B) Predicted structure of *T. pelophila* iCA, shaded by confidence (see legend). (C) Superposition of structures predicted for iCA homologs from *T. pelophila* and *T. aerophila* onto the empirically determined structure of iCA from *Nostoc* sp. PCC7120 [PDB ID 7C5V (29)], and a table of results from pairwise structural alignments of *Thiomicrospira* iCA proteins to iCA from *Nostoc* sp. PCC7120. Predicted structures for the *Thiomicrospira* iCA homologs were compared to 7C5V via TM-Align (44) implemented at RSCB Protein Data Bank (<https://www.rcsb.org/alignment>). Abbreviations: pLDDT, predicted local distance difference test, a measure of confidence in local structure (45); RMSD, root mean square deviation of distances in angstroms between alpha carbons in compared structures; TM-score, template modeling score, proportional to the agreement in topology between compared structures.

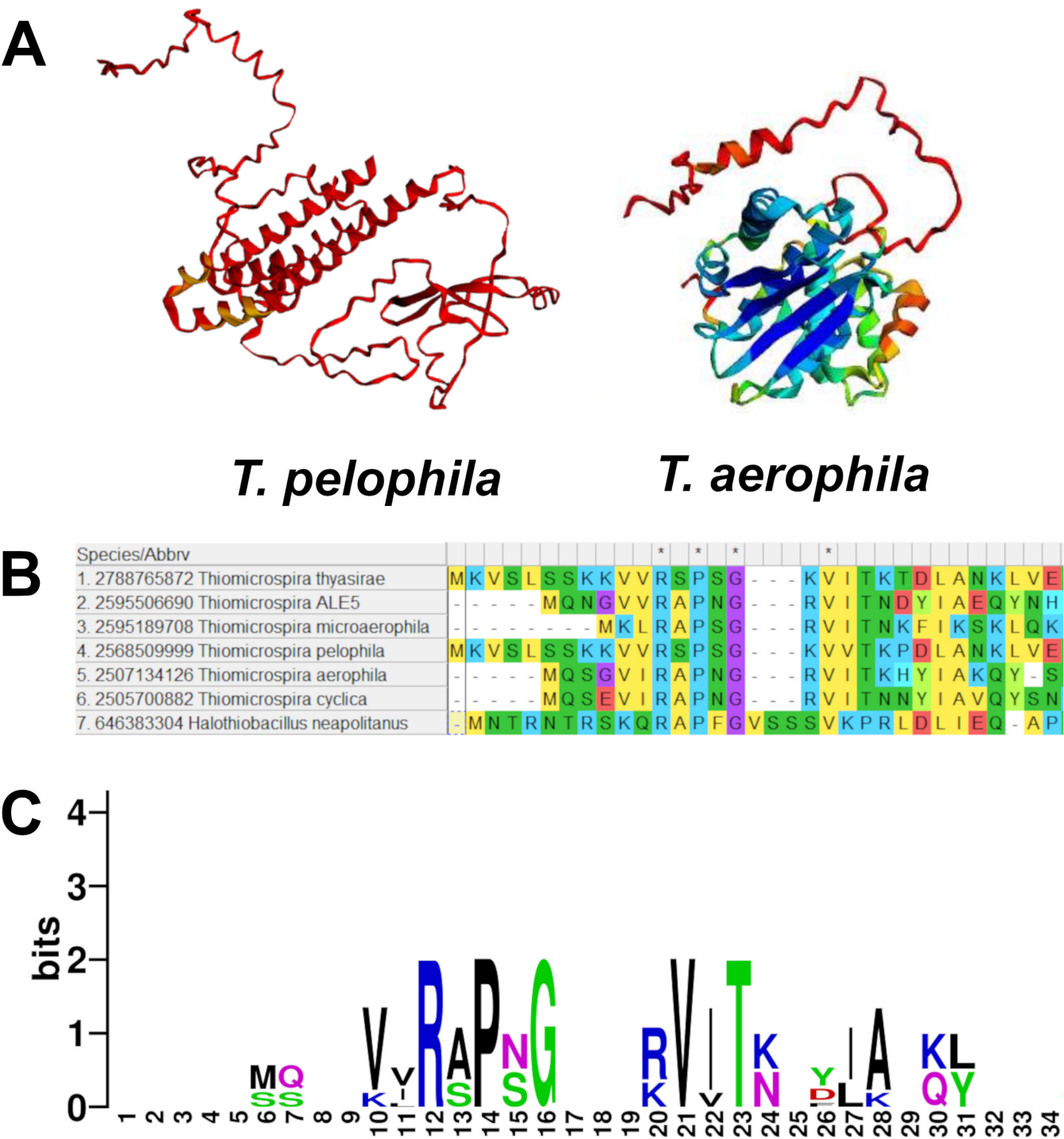


FIG 9 Comparison of CsoSX proteins from members of genus *Thiomicrospira*. (A) Structures predicted for CsoSX from *T. pelophila* and *T. aerophila* via AlphaFold2 (43), implemented via Colabfold (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) using default parameters, without comparison to PDB structures. Abbreviations: pLDDT, predicted local distance difference test, a measure of confidence in local structure (45). (B) Sequence alignment [via MUSCLE, implemented in MEGA11 (47, 48)], and (C) sequence logo created from the alignment of CsoSX amino acid sequences from members of genus *Thiomicrospira* with CsoSCA from *H. neapolitanus*. The complete logo is depicted in Fig. S6.

Despite this lack of predicted structural similarity to CsoSCA, it is likely that CsoSX is a modified version of the amino terminal region of CsoSCA. All CsoSX proteins share a short amino terminal sequence motif with CsoSCA proteins (RXPXGXXXV, corresponding

to *H. neapolitanus* CsoSCA residues 11–20; Fig. 9B and C). Additionally, the amino terminal region of CsoSX from *T. thyasirae* (residues 63–131) matches Interpro family IPR043065 [N-terminal domain of CsoSCA, $e < 2E-5$; <https://www.ebi.ac.uk/interpro/> (49)]. The remainder of the sequence (residues 132–332) has no match, which is unlike CsoSCA sequences, which also match IPR043066 (carboxysome shell carbonic anhydrase and C-terminal domain superfamily).

Conservation of the amino terminal residues of CsoSX suggests that the presence of the protein has been selected for, and CsoSX is indeed present in carboxysome preparations (Table 1; Table S1). However, CsoSX proteins appear to be unique to *Thiomicrospira*; none of the other CsoSX sequences from *Thiomicrospira* match Interpro families nor do they match sequences in the nonredundant protein database or structures available to searches with Foldseek [<https://search.foldseek.com/search> (50)]. The amino termini of CsoSCA proteins are disordered and interact with CbbL and CbbS [Fig. 10 (51)]. The amino termini of CsoSX proteins are predicted to lack much secondary structure; perhaps they too are disordered (Fig. 9). It is possible that CsoSX interacts with CbbL and CbbS and facilitates packing them into *Thiomicrospira* carboxysomes.

Redox sensitivity varies among ica proteins

The sensitivity of carboxysome CA activity to the reducing agent dithiothreitol (DTT) differs between *T. pelophila* and *aerophila* (Table 2). *T. pelophila* carboxysomal CA activity is insensitive to DTT, while that from *T. aerophila* is completely inhibited by this compound. Of the two ica proteins present in *T. aerophila* carboxysomes, ica115 is present at an order-of-magnitude excess relative to ica112 (Table 1), so this sensitivity to

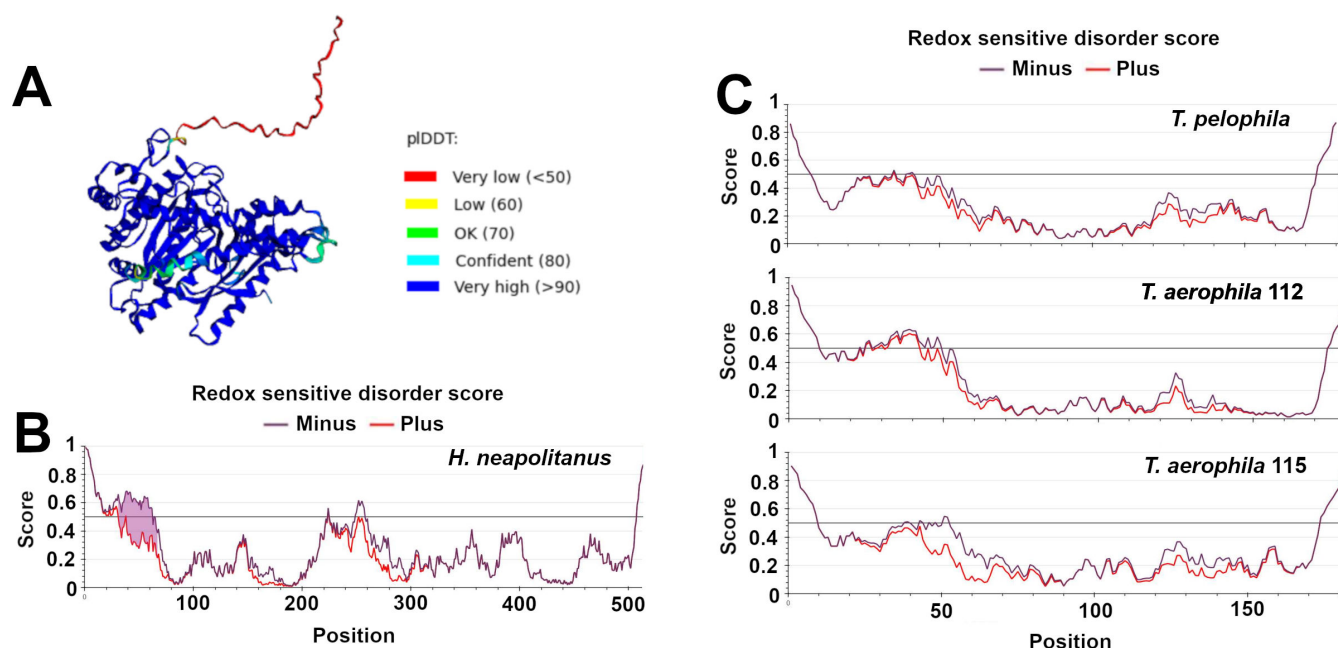


FIG 10 Potential redox sensitivity of amino termini of some carboxysome-associated carbonic anhydrase enzymes. (A) Structure of CsoSCA from *H. neapolitanus* predicted via AlphaFold2 (43), implemented via Colabfold (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>) using the “pdb100” option for template_mold since an empirically determined structure is available for this protein [PDB 2FGY (46)]. A predicted structure was generated because the published structure does not include the amino terminus. The published structure and this predicted structure are in good agreement (RMSD = 0.66; TM-score = 0.99; TM-Align) (44) and implemented at RSCB Protein Data Bank (<https://www.rcsb.org/alignment>). Predicted disordered regions of CsoSCA (B), and ica from members of genus *Thiomicrospira* (C), and their potential sensitivity to redox status. “Position” is with respect to the residue number in the amino acid sequence, and “score” is the disorder prediction score, which indicates the probability that a residue is located in a disordered region of the protein. “Minus” and “plus” refer to predicted disorder in the absence (minus; cysteine replaced by alanine) and presence (plus; native sequence) of redox-sensitive amino acids. Disordered regions and impact from redox status were predicted via IUPred2A [(52); <https://iupred2a.elte.hu/>]. Abbreviations: pLDDT, predicted local distance difference test, a measure of confidence in local structure (45).

DTT likely reflects an impact on iCA115 activity. Interestingly, CsoSCA from *H. neapolitanus* is also inhibited by DTT (39). For both *T. aerophila* iCA115 and *H. neapolitanus* CsoSCA, the relative disorder of the structure of the amino terminus is predicted to change due to redox status (e.g., due to disulfide bond formation between cysteine residues); such changes are not predicted for *T. pelophila* iCA or *T. aerophila* iCA112 (Fig. 10). The amino termini of the *T. aerophila* iCA115 and *H. neapolitanus* CsoSCA have more cysteine residues (two, C54 and C57; and three, C41, C49, and C65, respectively) than the others (one each; C54 for *T. pelophila* and C53 for *T. aerophila* iCA112). Perhaps the sensitivities of some iCA enzymes to DTT are due to the loss of intra- or intermolecular disulfide bonds in the presence of this reducing agent.

It is tempting to speculate that these α -carboxysomal CA enzymes are redox-regulated, as is the case for the β - and γ CA enzymes present in β -carboxysomes from *Cyanobacteria*, which are also inhibited by DTT (15, 16), and this redox sensitivity has been proposed to be beneficial during carboxysome assembly. Cells with CCMs have elevated concentrations of cytoplasmic HCO_3^- generated by active transport of DIC into the cell from the environment by membrane-spanning transporters (53). If carboxysomal CA were active in the cytoplasm during carboxysome assembly, it would convert cytoplasmic HCO_3^- into CO_2 , which would then be lost from the cell via diffusion through the cell membrane (54). It has been suggested that carboxysomal CA enzymes that are inactive in reducing environments, such as the cytoplasm, will not cause CO_2 leakage from cells during carboxysome assembly (16). Presumably, these enzymes are oxidized and active once packaged within carboxysomes.

Preventing CO_2 leakage during carboxysome assembly would seem to provide a strong selective advantage for carboxysomal inactivation by reducing conditions, which makes the distribution of redox sensitivity among iCA enzymes present in *Thiomicrospira* quite puzzling. If multiple cysteine residues at the amino terminus result in redox sensitivity, only three members of *Thiomicrospira* have redox-sensitive iCA. Like *T. aerophila*, both *T. cyclica* and *Thiomicrospira* sp. ALE5 have two genes encoding iCA, one of which (in each species) has two cysteine residues in its amino terminus and predicted redox sensitivity of structural disorder [via IUPred2A (52); <https://iupred2a.elte.hu/>] similar to iCA from *T. aerophila* (Fig. 10). All three of these species also have a gene encoding iCA without multiple amino terminal cysteine residues. The other four members of the genus only have iCA with amino termini having single cysteine residues. Based on the observed redox sensitivities of iCA from *T. pelophila* and *T. aerophila*, as well as inferences based on amino terminal sequences, only three members of *Thiomicrospira* have redox-sensitive iCA, and all seven species carry genes encoding iCA that is not sensitive to redox conditions. Perhaps the redox-insensitive version of this enzyme is an advantage for cells growing in highly reduced conditions, where hydrogen sulfide may be abundant enough to inactivate redox-sensitive iCA (55), or where activating the enzyme by oxidation is less thermodynamically favorable. This may explain why *T. microaerophila*, which requires low O_2 concentrations for growth (56), only has the redox-insensitive form (Fig. 7). It would be helpful to determine whether the amino terminal cysteines are indeed responsible for redox sensitivity in these iCA enzymes, and whether this redox sensitivity has a role in preventing CO_2 leakage during carboxysome assembly. It would also be helpful to understand how redox-insensitive carboxysomal CA enzymes are regulated during the assembly of these microcompartments.

iCA could offer a selective advantage in alkaline habitats

Thus far, *Thiomicrospira* is the only genus in which CsoSCA appears to have been replaced by iCA. The other genus in which iCA is common in the carboxysome locus is *Thioalkalivibrio*, though for these organisms, a *csoSCA* gene is also present in the locus (Fig. 1 and Fig. 7). Both of these genera consist entirely (*Thioalkalivibrio*) or mostly (*Thiomicrospira*) of alkaliphilic organisms with optimal growth at pH 9–10 (57, 58). Most have been isolated from soda lakes with pH values from 9.2 to 11 (57). Given that the solubility of Zn^{+2} and other metals is extremely low under alkaline conditions [e.g., the

solubility product of $\text{Zn}(\text{OH})_2$ is $\sim 2 \times 10^{-17}$ at 15°C – 40°C (59)], acquiring sufficient Zn^{+2} or other metals necessary for CA activity would be problematic. Some α CA enzymes are atypical in that they do not require metals for activity (29), and this trait may be common among them (30). A lack of dependence on metals would provide a big advantage for carboxysome function in these habitats and may explain why CsoSCA, which requires Zn^{+2} (39), was replaced by α CA in *Thiomicrospira*. For members of *Thioalkalivibrio*, it would be interesting to determine which form of CA is packed into carboxysomes (CsoSCA or α CA), and whether this depends on cultivation conditions (e.g., pH). It would be interesting to determine whether factors that favor α - and β CA in α -carboxysomes are similar to those driving the distribution of β - and γ CA in β -carboxysomes (12).

Given that autotrophic organisms grow in every imaginable habitat on Earth, it is not surprising that they have a diversity of biochemistries and physiological adaptations to facilitate CO_2 fixation at any pH, temperature, or solute composition amenable to life (5, 60). Carboxysome diversity is one such adaptation that could facilitate CO_2 fixation at a range of redox conditions or metal availabilities, and it is likely that more carboxysome diversity remains to be uncovered. Indeed, there are members of *Nitrosospora* (Phylum *Pseudomonadota*) and *Pseudonocardia* (Phylum *Actinomycetota*) whose carboxysome loci lack genes encoding any of the currently known forms of CA (20). An understanding of the full breadth of autotrophic adaptations to carbon fixation in all sorts of habitats will add to an understanding of life here and elsewhere. Furthermore, carboxysomes are being integrated into synthetic systems for generating compounds of industrial importance (61, 62); increasing the parameter space (e.g., pH) for the synthesis by native and engineered autotrophic microorganisms could be helpful for these efforts.

MATERIALS AND METHODS

Bacterial strains and cultivation conditions

T. pelophila DSM 1534^T was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ) and cultivated at 32°C under ambient air in thiosulfate supplemented artificial seawater medium [TASW, pH 7.5, $1 \mu\text{g L}^{-1}$ vitamin B12 (53, 63)]. *T. aerophila* DSM 13739^T was obtained from DSMZ and cultivated in DSMZ medium 925 (alkaliphilic sulfur respiring strains medium) and incubated at 32°C under ambient air.

Five liter bioreactors were used for large-scale cultivation (Eppendorf, BioFlo 120). The pH was monitored and maintained via the addition of 10 M KOH (pH 7.5 for *T. pelophila* and pH 9.5 for *T. aerophila*). O_2 electrodes monitored the oxygen concentration, and cultures were stirred and sparged with O_2 to maintain O_2 levels at 95% air saturation. For mutant strain *T. pelophila* α CA-135, which is unable to grow under ambient air, O_2 was maintained by sparging with 95% O_2 , 5% CO_2 (vol/vol).

E. coli Lemo21(DE3) $\Delta yadF \Delta cynT$ was used to identify genes encoding CA. With both genes encoding native CA (*yadF/cynT*) deleted, this strain cannot grow with air as the source of CO_2 (64). Heterologous expression of HCO_3^- transporters or CA enzymes restores its ability to grow under ambient air (64). *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ was cultivated in lysogeny broth [LB (65)], supplemented with the appropriate antibiotics (see below), under either an air or 5% CO_2 , 95% air (vol/vol) headspace (see below). Cultures were incubated at 37°C and 100 rpm.

Carboxysome purification

Volumes of 100–300 L of *T. pelophila* and *T. aerophila* culture were grown to early stationary phase in 5 or 10 L bioreactors (Eppendorf, BioFlo 120) and harvested via centrifugation (Sorvall SLA-1500 rotor, $3,795 \times g$, 10 min, 10°C – 20°C). Carboxysomes were purified from the cells via differential centrifugation and use of sucrose gradients (66). In order to rupture carboxysomes for enzyme assays, pellets were frozen for 30 min at -20°C and stored at -80°C . Carboxysome preparations were monitored via SDS-PAGE (67). LC-MS/MS was used to determine the identities of the proteins, as in reference (64).

Carboxysome preparations were also negatively stained with ammonium molybdate and examined with a FEI Morgani 268D transmission electron microscope (68).

CA assays

CA assays were conducted on purified carboxysomes and crude extracts from *E. coli* expressing genes from *T. pelophila* via stopped-flow spectrophotometry. To prepare for the assay, frozen carboxysomes were suspended in CA buffer (50 mM HEPES, 0.2 mM ZnSO_4 , and pH 8 with NaOH) via vortexing and incubated at room temperature for 30 min. Frozen cell pellets from 500 mL cultures of *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ heterologously expressing proteins from *T. pelophila* were thoroughly suspended in 2 mL CA buffer supplemented with lysozyme and bovine pancreas deoxyribonuclease I (Thermo Scientific, 0.1 mg each). Two milliliters of BPER-II Bacterial Protein Extraction Reagent (Thermo Scientific) was added to facilitate cell lysis, and samples were sonicated four times for 30 s while incubated in ice water. After sonication, samples were incubated at room temperature for 15 min while gyrated at 100 rpm (C10 platform shaker, New Brunswick Scientific).

Stopped-flow spectrophotometric CA assays were conducted similarly to references (39, 69), with an SX20 spectrophotometer (Applied Photophysics), using phenol red to track the fall in pH caused by CO_2 conversion to H_2CO_3 (and deprotonation to form $\text{HCO}_3^- + \text{H}^+$). Absorbance values (A_{557}) taken 200 times per second, beginning at 0.1 s and ending at 0.2 s, were used to calculate the initial rate of CO_2 hydration since measurements prior to 0.1 s were erratic due to transients resulting from sample injection into the reaction chamber, and those after 0.2 s demonstrated a slow decrease in the rate of net CO_2 hydration, likely due to back-reaction (H_2CO_3 dehydration). A similar diminishment of apparent rates of CO_2 hydration by bovine carbonic anhydrase (Sigma-Aldrich, C3934) was observed in pilot studies to optimize the CA assay for the study presented here and is also apparent in published CA assays of carboxysomes from *H. neapolitanus* [as measured by pH indicator protonation (68)]. Absorbance values from these timepoints were stripped from the individual Excel files from each run and combined into a single file using a Python script (70) available at Github (<https://github.com/scooterboi85/CSV-Conerter>). Changes in A_{557} were converted to changes in $[\text{H}^+]$ by using ΔA_{557} values from titrating 1:1 mixtures of CA buffer and indicator solution with a standardized 6 N HCl solution (Thermo Fisher), which provided a measured change of 15.69 mM H^+ per absorbance unit at 557 nm. Changes in $[\text{H}^+]$ were converted to changes in CO_2 concentrations by assuming 1 H^+ produced per CO_2 hydrated ($\text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{H}_2\text{CO}_3 \rightarrow \text{H}^+ + \text{HCO}_3^-$). Each sample was run in triplicate, and the rates of CO_2 hydration were compared to those measured for identical samples that had been denatured by incubating them at 95°C for 30 min (also run in triplicate) to determine whether CO_2 hydration by the sample contents was primarily chemical or potentially catalyzed by CA. For those samples whose rate of CO_2 hydration significantly exceeded those measured in denatured samples (see Statistical analyses, below), CA activities were calculated by subtracting the rates from boiled samples from the rates of the unboiled samples.

Random mutagenesis of *T. pelophila* and characterization of mutant phenotypes

Random mutagenesis of *T. pelophila* was undertaken as in reference (71). In short, a clone library of 10,000 mutant strains of *T. pelophila* was generated by mating with *E. coli* BW20767 carrying plasmid pRL-27, which encodes transposon Tn5 (72). Once transformed into *T. pelophila*, Tn5 inserts randomly into the genome, generating a library of mutant strains, selectable by kanamycin since Tn5 encodes resistance to this antibiotic. The library was maintained on a solid medium containing 25 mg L^{-1} kanamycin under a high CO_2 headspace [5% CO_2 /95% air (vol/vol)]. This library was then screened for strains that had lost their ability to grow with air as the source of CO_2 (~0.04% CO_2) using high

throughput 96-well growth assays (71). CO₂-sensitive strains were set aside for further characterization.

Characterization of mutant phenotype began by verifying CO₂ sensitivity by growing each of these strains individually in 20 mL TASW in paired cultures (air, 5% CO₂/95% air) and monitoring growth via optical density at 480 nm. Transposon insertion sites were identified using arbitrary PCR (71) to amplify the gDNA adjacent to Tn5 insertion for sequencing (Psomagen, Inc.). CO₂-sensitive mutants were also examined for carboxysome presence via transmission electron microscopy, as in reference (22).

Heterologous expression of genes of interest in *E. coli*

To clarify if genes interrupted in CO₂-sensitive *T. pelophila* mutant strains encode CA, these genes were cloned into the plasmid pBAD, for expression in *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ under control of the *araBp* promoter. Plasmids carrying either *csoS2* or *csoSX* were synthesized by the Department of Energy Joint Genome Institute. Plasmids carrying the gene encoding iCA were synthesized by amplifying this gene from *T. pelophila* via PCR and inserting it into pBAD (Thermo-Fisher Scientific). Plasmids were initially propagated in *E. coli* Top10 in lysogeny medium supplemented with 100 mg L⁻¹ ampicillin and subsequently transformed into chemically competent *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ (64), which were cultivated under a high CO₂ headspace on lysogeny media supplemented with antibiotics (30 mg L⁻¹ apramycin and 30 mg L⁻¹ chloramphenicol for expressing CsoSX, CsoS2; 100 mg L⁻¹ ampicillin and 30 mg L⁻¹ chloramphenicol for expressing iCA).

Inocula for the growth assay for CA activity were prepared by adding cells to 20 mL of LB supplemented with the appropriate antibiotics and cultivating overnight under 5% CO₂, 95% air (vol/vol). The following morning, inocula of 100 μ L were introduced into 100 mL cultures containing 6 mM arabinose (induced), 11 mM glucose (repressed), or no additions (neutral) (73). Three replicates were run per condition. Additionally, *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$, which was not carrying a plasmid (no plasmid control), was run alongside each experiment. Flasks were incubated with air headspace at 37°C, 100 rpm, and growth was monitored via spectrophotometry at OD₆₀₀ for 48 h.

Larger-scale cultures were used to prepare biomass for CA assays of *E. coli* (Lemo21(DE3) $\Delta yadF \Delta cynT$ expressing iCA from *T. pelophila*). Five hundred milliliters LB (supplemented with 100 mg L⁻¹ ampicillin and 30 mg L⁻¹ chloramphenicol) was grown overnight, at 37°C, 100 rpm, under 5% CO₂, 95% air (vol/vol) headspace. The next day, the culture was harvested via centrifugation (10 min, 3,795 $\times g$, Sorvall SLA-1500, 10°C–20°C). The cell pellet was re-suspended in 2.5 L of LB containing arabinose to induce expression. The 2.5 L culture was split into 0.5 L portions, which were incubated at 37°C, 100 rpm under ambient air, and cell growth (OD₆₀₀) was monitored. Five hundred milliliter portions were harvested via centrifugation after 0, 4, 6, 8, and 24 h. Cell pellets were stored at –80°C, until assayed for CA activity. The successful expression of *T. pelophila* genes by *E. coli* (Lemo21(DE3) $\Delta yadF \Delta cynT$) was verified by subjecting proteins extracted from cell pellets to LC-MS/MS analysis as described above for carboxysome preparations.

Phylogenetic analyses

Homologs to the *Thiomicrospira* iCA genes were collected from the Integrated Microbial Genomes database [IMG, <https://img.jgi.doe.gov/> (41)] via BLASTp (74), with iCA homologs from *Thiomicrospira* as query sequences. Sequences from reference (30) were gathered from the nonredundant protein sequences database (<https://www.ncbi.nlm.nih.gov/>). Amino acid sequences were aligned using MUSCLE (47), and phylogenetic trees were constructed using the maximum likelihood method (75). Alignment, model evaluation, and phylogenetic analysis were all implemented in MEGA 11 (48).

Statistical analyses

CO₂ hydration rates by carboxysomes were compared to rates measured for assay buffer, boiled carboxysomes, and carboxysomes in the presence of potential inhibitors or solvent (*n* = 3 for each condition). For carboxysome preparations from each strain (*T. pelophila*, *T. aerophila*, or *T. pelophila* iCA-135), rates were subjected to ANOVA, with a post-hoc Tukey's test for significance of difference among the treatments. CO₂ hydration rates by cell lysate from *E. coli* Lemo21(DE3) $\Delta yadF \Delta cynT$ expressing genes from *T. pelophila* were analyzed by comparing boiled (*n* = 3) and unboiled (*n* = 3) rates via Welch *t* tests. All statistical analyses were implemented in R v4.3.2 using the R Stats Package (76).

ACKNOWLEDGMENTS

We are grateful for support from the National Science Foundation (NSF-MCB-1952676 to K.M.S.).

We are also grateful to C. Kerfeld, M. Sutter, G. Cannon, and S. Heinhorst for helpful discussions of this work with us during manuscript preparation and to Carolina Sarmiento Bravo for assistance on the figures. We are also grateful to the anonymous reviewers for their helpful suggestions, which improved the manuscript.

J.W.: formal analysis, investigation, methodology, writing—original draft, and writing—review and editing. D.F., T.M., A.J.M., R.R.P., D.W., and Y.X.Z.: investigation, and writing—review and editing. A.G.: investigation, software, and writing—review and editing. M.C., J.C., J.K., R.K., C.L., N.L., T.M., K.T.B., T.T., C.D., and D.C.: investigation, and writing—review and editing. K.M.S.: conceptualization, formal analysis, funding acquisition, investigation, methodology, supervision, visualization, writing—original draft, and writing—review and editing.

AUTHOR AFFILIATION

¹Integrative Biology Department, University of South Florida, Tampa, Florida, USA

AUTHOR ORCID*s*

Kathleen M. Scott  <http://orcid.org/0000-0002-9407-518X>

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation (NSF)	NSF-MCB-1952676	Kathleen M. Scott

AUTHOR CONTRIBUTIONS

Jana Wiescholke, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing | Daniella Fuller, Investigation, Writing – review and editing | Arin Gahramanova, Investigation, Software, Writing – review and editing | Terrence Millen, Investigation, Writing – review and editing | Ashianna J. Mislav, Investigation, Writing – review and editing | Ren R. Payne, Investigation, Writing – review and editing | Daniel P. Walsh, Investigation, Writing – review and editing | YuXuan Zhao, Investigation, Writing – review and editing | Madilyn Carney, Investigation, Writing – review and editing | Jaden Cross, Investigation, Writing – review and editing | John Kashem, Investigation, Writing – review and editing | Ruchi Korde, Investigation, Writing – review and editing | Christine Lacy, Investigation, Writing – review and editing | Noah Lyons, Investigation, Writing – review and editing | Tori Mason, Investigation, Writing – review and editing | Kayla Torres-Betancourt, Investigation, Writing – review and editing | Tyler Trapnell, Investigation, Writing – review and editing | Clare L. Dennison, Investigation, Writing – review and editing | Dale Chaput, Investigation, Writing – review and editing | Kathleen M. Scott, Conceptualization, Formal analysis, Funding acquisition, Investigation,

Methodology, Supervision, Visualization, Writing – original draft, Writing – review and editing

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Supplemental material (AEM01075-24-S0001.pdf). Tables S1 and S2; Figures S1 to S6.

REFERENCES

- Partensky F, Blanchot J, Vaulot D. 1999. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. *Bull Inst Oceanogr (Monaco)* 0:457–475.
- Kerfeld CA, Aussignargues C, Zarzycki J, Cai F, Sutter M. 2018. Bacterial microcompartments. *Nat Rev Microbiol* 16:277–290. <https://doi.org/10.1038/nrmicro.2018.10>
- Hagemann M, Kern R, Maurino VG, Hanson DT, Weber APM, Sage RF, Bauwe H. 2016. Evolution of photorespiration from cyanobacteria to land plants, considering protein phylogenies and acquisition of carbon concentrating mechanisms. *J Exp Bot* 67:2963–2976. <https://doi.org/10.1093/jxb/erw063>
- Badger MR, Hanson D, Price GD. 2002. Evolution and diversity of CO₂ concentrating mechanisms in cyanobacteria. *Funct Plant Biol* 29:161–173. <https://doi.org/10.1071/PP01213>
- Scott KM, Payne RR, Gahramanova A. 2024. Widespread dissolved inorganic carbon-modifying toolkits in genomes of autotrophic bacteria and archaea and how they are likely to bridge supply from the environment to demand by autotrophic pathways. *Appl Environ Microbiol* 90:e0155723. <https://doi.org/10.1128/aem.01557-23>
- Cooper TG, Filmer D, Wishnick M, Lane MD. 1969. The active species of “CO₂” utilized by ribulose diphosphate carboxylase. *J Biol Chem* 244:1081–1083.
- Shively JM, Ball F, Brown DH, Saunders RE. 1973. Functional organelles in prokaryotes: polyhedral inclusions (carboxysomes) of *Thiobacillus neapolitanus*. *Science* 182:584–586. <https://doi.org/10.1126/science.182.4112.584>
- So A-C, Espie GS, Williams EB, Shively JM, Heinhorst S, Cannon GC. 2004. A novel evolutionary lineage of carbonic anhydrase (epsilon class) is a component of the carboxysome shell. *J Bacteriol* 186:623–630. <https://doi.org/10.1128/JB.186.3.623-630.2004>
- Kerfeld CA, Melnicki MR. 2016. Assembly, function and evolution of cyanobacterial carboxysomes. *Curr Opin Plant Biol* 31:66–75. <https://doi.org/10.1016/j.pbi.2016.03.009>
- Cameron JC, Sutter M, Kerfeld CA. 2014. The carboxysome: function, structure and cellular dynamics. In Flores E, Herrero A (ed), *The cell biology of cyanobacteria*. Caister Academic Press, UK.
- Badger MR, Bek EJ. 2008. Multiple RubisCO forms in proteobacteria: their functional significance in relation to CO₂ acquisition by the CBB cycle. *J Exp Bot* 59:1525–1541. <https://doi.org/10.1093/jxb/ern297>
- Rae BD, Long BM, Badger MR, Price GD. 2013. Functions, compositions, and evolution of the two types of carboxysomes: polyhedral microcompartments that facilitate CO₂ fixation in cyanobacteria and some proteobacteria. *Microbiol Mol Biol Rev* 77:357–379. <https://doi.org/10.1128/MMBR.00061-12>
- Sommer M, Cai F, Melnicki MR, Kerfeld CA. 2017. Beta carboxysome bioinformatics: identification, evolution and implications of new classes of shell protein genes and core locus constraints. *J Exper Bot* 68:3841–3855. <https://doi.org/10.1093/jxb/erx115>
- Yu JW, Price GD, Song L, Badger MR. 1992. Isolation of a putative carboxysomal carbonic anhydrase gene from the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol* 100:794–800. <https://doi.org/10.1104/pp.100.2.794>
- Price GD, Coleman JR, Badger MR. 1992. Association of carbonic anhydrase activity with carboxysomes isolated from the cyanobacterium *Synechococcus* PCC7942. *Plant Physiol* 100:784–793. <https://doi.org/10.1104/pp.100.2.784>
- Peña KL, Castel SE, de Araujo C, Espie GS, Kimber MS. 2010. Structural basis of the oxidative activation of the carboxysomal gamma-carbonic anhydrase, CcmM. *Proc Natl Acad Sci U S A* 107:2455–2460. <https://doi.org/10.1073/pnas.0910866107>
- de Araujo C, Arefeen D, Tadesse Y, Long BM, Price GD, Rowlett RS, Kimber MS, Espie GS. 2014. Identification and characterization of a carboxysomal γ-carbonic anhydrase from the cyanobacterium *Nostoc* sp. PCC 7120. *Photosynth Res* 121:135–150. <https://doi.org/10.1007/s11120-014-0018-4>
- Matsuda Y, Nawaly H, Yoneda K. 2022. Carbonic anhydrase, p 167–195. In *Blue planet, red and green photosynthesis*
- Scott KM, Leonard J, Boden R, Chaput D, Dennison C, Haller E, Harmer TL, Anderson A, Arnold T, Brand J, et al. 2019. Diversity in CO₂ concentrating mechanisms among chemolithoautotrophs from the genera *Hydrogenovibrio*, *Thiomicrothabodus*, and *Thiomicrospira*, ubiquitous in sulfidic habitats worldwide. *Appl Environ Microbiol* 85:e02096–18. <https://doi.org/10.1128/AEM.02096-18>
- Sutter M, Kerfeld CA, Scott KM, USF Genomics Class 2020, USF Genomics Class 2021. 2021. Atypical carboxysome loci: JEEPs or Junk? *Front Microbiol* 13. <https://doi.org/10.3389/fmicb.2022.872708>
- Scott KM, Williams J, Porter CMB, Russel S, Harmer TL, Paul JH, Antonen KM, Bridges MK, Camper GJ, Campla CK, et al. 2018. Genomes of ubiquitous marine and hypersaline *Hydrogenovibrio*, *Thiomicrothabodus* and *Thiomicrospira* spp. encode a diversity of mechanisms to sustain chemolithoautotrophy in heterogeneous environments. *Environ Microbiol* 20:2686–2708. <https://doi.org/10.1111/1462-2920.14090>
- Dobrinski KP, Enkemann SA, Yoder SJ, Haller E, Scott KM. 2012. Transcriptional response of the sulfur chemolithoautotroph *Thiomicrospira crunogena* to dissolved inorganic carbon limitation. *J Bacteriol* 194:2074–2081. <https://doi.org/10.1128/JB.06504-11>
- Igarashi Y, Kodama T. 1996. Genes related to carbon dioxide fixation in *Hydrogenovibrio marinus* and *Pseudomonas hydrothermophila*, p 88–93. In Lidstrom ME, Tabita FR (ed), *Microbial growth on C1 compounds*. Kluwer Academic Publishers.
- Kuenen JG, Veldkamp H. 1972. *Thiomicrospira pelophila*, gen. n., sp. n., a new obligately chemolithotrophic colourless sulfur bacterium. *Antonie Van Leeuwenhoek* 38:241–256. <https://doi.org/10.1007/BF02328096>
- Sorokin DY, Lysenko AM, Mityushina LL, Tourova TP, Jones BE, Rainey FA, Robertson LA, Kuenen GJ. 2001. *Thioalkalimicrobium aerophilum* gen. nov., sp. nov. and *Thioalkalimicrobium sibericum* sp. nov., and *Thioalkalivibrio versutus* gen. nov., sp. nov., *Thioalkalivibrio nitratis* sp. nov., novel and *Thioalkalivibrio denitrificans* sp. nov., novel obligately alkaliphilic and obligately chemolithoautotrophic sulfur-oxidizing bacteria from soda lakes. *Int J Syst Evol Microbiol* 51:565–580. <https://doi.org/10.1099/00207713-51-2-565>
- Sorokin DY, Gorlenko VM, Tourova TP, Tsapin AI, Nealson KH, Kuenen GJ. 2002. *Thioalkalimicrobium cyclicum* sp. nov. and *Thioalkalivibrio jannaschii* sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing bacteria from hypersaline alkaline mono lake (California). *Int J Syst Evol Microbiol* 52:913–920. <https://doi.org/10.1099/00207713-52-3-913>
- Jensen EL, Clement R, Kosta A, Maberly SC, Gontero B. 2019. A new widespread subclass of carbonic anhydrase in marine phytoplankton. *ISME J* 13:2094–2106. <https://doi.org/10.1038/s41396-019-0426-8>
- Del Prete S, Nocentini A, Supuran CT, Capasso C. 2020. Bacterial γ-carbonic anhydrase: a new active class of carbonic anhydrase identified

- in the genome of the gram-negative bacterium *Burkholderia territorii*. *J Enzyme Inhib Med Chem* 35:1060–1068. <https://doi.org/10.1080/14756366.2020.1755852>
29. Hirakawa Y, Senda M, Fukuda K, Yu HY, Ishida M, Taira M, Kinbara K, Senda T. 2021. Characterization of a novel type of carbonic anhydrase that acts without metal cofactors. *BMC Biol* 19:105. <https://doi.org/10.1186/s12915-021-01039-8>
 30. Nocentini A, Supuran CT, Capasso C. 2021. An overview on the recently discovered iota-carbonic anhydrases. *J Enzyme Inhib Med Chem* 36:1988–1995. <https://doi.org/10.1080/14756366.2021.1972995>
 31. Tsai Y-C, Liew L, Guo Z, Liu D, Mueller-Cajar O. 2022. The CbbQO-type rubisco activases encoded in carboxysome gene clusters can activate carboxysomal form IA rubiscos. *J Biol Chem* 298:101476. <https://doi.org/10.1016/j.jbc.2021.101476>
 32. Paoli GC, Vichivanives P, Tabita FR. 1998. Physiological control and regulation of the *Rhodobacter capsulatus* cbb operons. *J Bacteriol* 180:4258–4269. <https://doi.org/10.1128/JB.180.16.4258-4269.1998>
 33. Shively JM, Bradburne CE, Aldrich HC, Bobik TA, Mehlmán JL, Jin S, Baker SH. 1998. Sequence homologs of the carboxysomal polypeptide CsoS1 of the thiobacilli are present in cyanobacteria and enteric bacteria that form carboxysomes - polyhedral bodies. *Can J Bot* 76:906–916. <https://doi.org/10.1139/cjb-76-6-906>
 34. Tanaka S, Kerfeld CA, Sawaya MR, Cai F, Heinhorst S, Cannon GC, Yeates TO. 2008. Atomic-level models of the bacterial carboxysome shell. *Science* 319:1083–1086. <https://doi.org/10.1126/science.1151458>
 35. Klein MG, Zwart P, Bagby SC, Cai F, Chisholm SW, Heinhorst S, Cannon GC, Kerfeld CA. 2009. Identification and structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *J Mol Biol* 392:319–333. <https://doi.org/10.1016/j.jmb.2009.03.056>
 36. Cai F, Dou Z, Bernstein SL, Leverenz R, Williams EB, Heinhorst S, Shively J, Cannon GC, Kerfeld CA. 2015. Advances in understanding carboxysome assembly in *Prochlorococcus* and *Synechococcus* implicate CsoS2 as a critical component. *Life (Basel, Switzerland)* 5:1141–1171. <https://doi.org/10.3390/life5021141>
 37. Wheatley NM, Eden KD, Ngo J, Rosinski JS, Sawaya MR, Cascio D, Collazo M, Hoveida H, Hubbell WL, Yeates TO. 2016. A PII-like protein regulated by bicarbonate: structural and biochemical studies of the carboxysome-associated CPII protein. *J Mol Biol* 428:4013–4030. <https://doi.org/10.1016/j.jmb.2016.07.015>
 38. Wheatley NM, Sundberg CD, Gidaniyan SD, Cascio D, Yeates TO. 2014. Structure and identification of a pterin dehydratase-like protein as a ribulose-bisphosphate carboxylase/oxygenase (RuBisCO) assembly factor in the α -carboxysome. *J Biol Chem* 289:7973–7981. <https://doi.org/10.1074/jbc.M113.531236>
 39. Heinhorst S, Williams EB, Cai F, Murin CD, Shively JM, Cannon GC. 2006. Characterization of the carboxysomal carbonic anhydrase CsoSCA from *Halothiobacillus neapolitanus*. *J Bacteriol* 188:8087–8094. <https://doi.org/10.1128/JB.00990-06>
 40. Jensen EL, Receveur-Brechot V, Hachemane M, Wils L, Barbier P, Parsiegla G, Gontero B, Launay H. 2021. Structural contour map of the iota carbonic anhydrase from the diatom *Thalassiosira pseudonana* using a multiprong approach. *Int J Mol Sci* 22:8723. <https://doi.org/10.3390/ijms22168723>
 41. Chen I-M, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, Hajek P, Ritter SJ, Webb C, Wu D, Varghese NJ, Reddy TBK, Mukherjee S, Ovchinnikova G, Nolan M, Seshadri R, Roux S, Visel A, Woyke T, Eloe-Fadros EA, Kyrpides NC, Ivanova NN. 2023. The IMG/M data management and analysis system v.7: content updates and new features. *Nucleic Acids Res* 51:D723–D732. <https://doi.org/10.1093/nar/gkac976>
 42. Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. *Mol Biol Evol* 25:1307–1320. <https://doi.org/10.1093/molbev/msn067>
 43. Mirdita M, Schütze K, Moriawaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. *Nat Methods* 19:679–682. <https://doi.org/10.1038/s41592-022-01488-1>
 44. Zhang Y, Skolnick J. 2005. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33:2302–2309. <https://doi.org/10.1093/nar/gki524>
 45. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with alphaFold. *Nature* 596:583–589. <https://doi.org/10.1038/s41586-021-03819-2>
 46. Sawaya MR, Cannon GC, Heinhorst S, Tanaka S, Williams EB, Yeates TO, Kerfeld CA. 2006. The structure of beta-carbonic anhydrase from the carboxysomal shell reveals a distinct subclass with one active site for the price of two. *J Biol Chem* 281:7546–7555. <https://doi.org/10.1074/jbc.M510464200>
 47. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>
 48. Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 38:3022–3027. <https://doi.org/10.1093/molbev/msab120>
 49. Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L, et al. 2023. InterPro in 2022. *Nucleic Acids Res* 51:D418–D427. <https://doi.org/10.1093/nar/gkac993>
 50. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M. 2024. Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* 42:243–246. <https://doi.org/10.1038/s41587-023-01773-0>
 51. Blikstad C, Dugan EJ, Laughlin TG, Turnšek JB, Liu MD, Shoemaker SR, Vogiatzi N, Remis JP, Savage DF. 2023. Identification of a carbonic anhydrase-Rubisco complex within the alpha-carboxysome. *Proc Natl Acad Sci U S A* 120:e2308600120. <https://doi.org/10.1073/pnas.2308600120>
 52. Mészáros B, Erdos G, Dosztányi Z. 2018. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res* 46:W329–W337. <https://doi.org/10.1093/nar/gky384>
 53. Dobrinski KP, Longo DL, Scott KM. 2005. The carbon-concentrating mechanism of the hydrothermal vent chemolithoautotroph *Thiomicrospira crunogena*. *J Bacteriol* 187:5761–5766. <https://doi.org/10.1128/JB.187.16.5761-5766.2005>
 54. Price GD, Badger MR. 1989. Expression of human carbonic anhydrase in the cyanobacterium *Synechococcus* PCC7942 creates a high CO₂-requiring phenotype. *Plant Physiol* 91:505–513. <https://doi.org/10.1104/pp.91.2.505>
 55. Cuevasanta E, Lange M, Bonanata J, Coitiño EL, Ferrer-Sueta G, Filipovic MR, Alvarez B. 2015. Reaction of hydrogen sulfide with disulfide and sulfenic acid to form the strongly nucleophilic persulfide. *J Biol Chem* 290:26866–26880. <https://doi.org/10.1074/jbc.M115.672816>
 56. Sorokin DY, Foti M, Pinkart HC, Muyzer G. 2007. Sulfur-oxidizing bacteria in soap lake (Washington State), a meromictic, haloalkaline lake with an unprecedented high sulfide content. *Appl Environ Microbiol* 73:451–455. <https://doi.org/10.1128/AEM.02087-06>
 57. Sorokin DY, Kuenen JG, Muyzer G. 2011. The microbial sulfur cycle at extremely haloalkaline conditions of soda lakes. *Front Microbiol* 2:44. <https://doi.org/10.3389/fmicb.2011.00044>
 58. Boden R, Scott KM, Williams J, Russel S, Antonen K, Rae AW, Hutt LP. 2017. An evaluation of thiomicrospira, hydrogenovibrio and *Thioalkalimicrobium*: reclassification of four species of *Thiomicrospira* to each *Thiomicrospira* gen. nov. and hydrogenovibrio, and reclassification of all four species of *Thioalkalimicrobium* to *Thiomicrospira*. *Int J Syst Evol Micro* 67:1140–1151.
 59. Reichle RA, McCurdy KG, Hepler LG. 1975. Zinc hydroxide: solubility product and hydroxy-complex stability constants from 12.5–75 °C. *Can J Chem* 53:3841–3845. <https://doi.org/10.1139/v75-556>
 60. Berg IA. 2011. Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl Environ Microbiol* 77:1925–1936. <https://doi.org/10.1128/AEM.02473-10>
 61. Bonacci W, Teng PK, Afonso B, Niederholtmeyer H, Grob P, Silver PA, Savage DF. 2012. Modularity of a carbon-fixing protein organelle. *Proc Natl Acad Sci U S A* 109:478–483. <https://doi.org/10.1073/pnas.1108557109>
 62. Ferlez B, Sutter M, Kerfeld CA. 2019. A designed bacterial microcompartment shell with tunable composition and precision cargo loading. *Metab Eng* 54:286–291. <https://doi.org/10.1016/j.mbs.2019.04.011>
 63. Jannasch HW, Wirsen CO, Nelson DC, Robertson LA. 1985. *Thiomicrospira crunogena* sp. nov., a colorless, sulfur-oxidizing bacterium from a deep-sea hydrothermal vent. *Int J Syst Bacteriol* 35:422–424. <https://doi.org/10.1099/00207713-35-4-422>

64. Schmid S, Chaput D, Breitbart M, Hines R, Williams S, Gossett HK, Parsi SD, Peterson R, Whittaker RA, Tarver A, Scott KM. 2021. Dissolved inorganic carbon-accumulating complexes from autotrophic bacteria from extreme environments. *J Bacteriol* 203:e0037721. <https://doi.org/10.1128/JB.00377-21>
65. Bertani G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* 62:293–300. <https://doi.org/10.1128/jb.62.3.293-300.1951>
66. Menning KJ, Menon BB, Fox G, Scott KM, USF MCB4404L. 2012. 2016. Dissolved inorganic carbon uptake in *Thiomicrospira crunogena* XCL-2 is Δp - and ATP-sensitive and enhances RubisCO-mediated carbon fixation. *Arch Microbiol* 198:149–159. <https://doi.org/10.1007/s00203-015-1172-6>
67. Sambrook J, Russell DW. 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
68. Dou Z, Heinhorst S, Williams EB, Murin CD, Shively JM, Cannon GC. 2008. CO₂ fixation kinetics of *Halothiobacillus neapolitanus* mutant carboxysomes lacking carbonic anhydrase suggest the shell acts as a diffusional barrier for CO₂. *J Biol Chem* 283:10377–10384. <https://doi.org/10.1074/jbc.M709285200>
69. Khalifah RG. 1971. The carbon dioxide hydration activity of carbonic anhydrase: I. Stop-flow kinetic studies on the native human isoenzymes B and C. *J Biol Chem* 246:2561–2573. [https://doi.org/10.1016/S0021-9258\(18\)62326-9](https://doi.org/10.1016/S0021-9258(18)62326-9)
70. Rossum Gv. 1995. Python language reference, v3.12. Python Software Foundation, Amsterdam.
71. Mangiapia M, Brown T-RW, Chaput D, Haller E, Harmer TL, Hashemy Z, Keeley R, Leonard J, Mancera P, Nicholson D, Stevens S, Wanjugi P, Zabinski T, Pan C, Scott KM, USF MCB4404L. 2017. Proteomic and mutant analysis of the CO₂ concentrating mechanism of hydrothermal vent chemolithoautotroph *Thiomicrospira crunogena*. *J Bacteriol* 199:e00871-16. <https://doi.org/10.1128/JB.00871-16>
72. Larsen RA, Wilson MM, Guss AM, Metcalf WW. 2002. Genetic analysis of pigment biosynthesis in *Xanthobacter autotrophicus* Py2 using a new, highly efficient transposon mutagenesis system that is functional in a wide variety of bacteria. *Arch Microbiol* 178:193–201. <https://doi.org/10.1007/s00203-002-0442-2>
73. Hendrickson W, Schleif RF. 1984. Regulation of the *Escherichia coli* L-arabinose operon studied by gel electrophoresis DNA binding assay. *J Mol Biol* 178:611–628. [https://doi.org/10.1016/0022-2836\(84\)90241-9](https://doi.org/10.1016/0022-2836(84)90241-9)
74. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>
75. Felsenstein J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368–376. <https://doi.org/10.1007/BF01734359>
76. R Core Team. 2023. R foundation for statistical computing. Vienna, Austria.