



This is a repository copy of *Structural basis and evolutionary pathways of glycerol-1-phosphate transport in marine bacteria*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/235578/>

Version: Published Version

---

**Article:**

Wang, N., Westermann, L.M. orcid.org/0000-0003-3479-4526, Li, M. et al. (9 more authors) (2025) Structural basis and evolutionary pathways of glycerol-1-phosphate transport in marine bacteria. *Proceedings of the National Academy of Sciences*, 122 (50). e2524546122. ISSN: 0027-8424

<https://doi.org/10.1073/pnas.2524546122>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>



# Structural basis and evolutionary pathways of glycerol-1-phosphate transport in marine bacteria

Ning Wang<sup>a,b,1</sup>, Linda M. Westermann<sup>c,1</sup>, Mingyu Li<sup>c,1</sup>, Chun-Yang Li<sup>a,1</sup>, Andrew R. J. Murphy<sup>c</sup>, Zengtian Gu<sup>b</sup>, Eleonora Silvano<sup>c,d</sup>, Claudia A. Blindauer<sup>e</sup>, Ian D. E. A. Lidbury<sup>f</sup>, Yu-Zhong Zhang<sup>a,b,2</sup>, David J. Scanlan<sup>c,2</sup>, and Yin Chen<sup>d,g,2</sup>

Affiliations are included on p. 9.

Edited by Edward DeLong, University of Hawai'i at Manoa, Honolulu, HI; received September 3, 2025; accepted November 13, 2025

All cells use lipid membranes to maintain cellular integrity and function, though Archaea utilize lipids composed of glycerol-1-phosphate (G1P), while Bacteria and Eukaryotes use glycerol-3-phosphate (G3P). Given that Archaea contribute significantly to global marine biomass, accounting for 0.3 gigatonnes (Gt) of carbon in the oceans, we aimed to uncover how archaeal G1P is recycled by marine microorganisms. Through a multi-disciplinary approach combining microbiology, biochemistry, and structural biology, we identified a G1P transporter in marine bacteria, which we named GpxB. Phylogenetic analysis revealed that GpxB belongs to the organic phosphonate transporter (PhnT) family and is widely distributed in the marine microbiome, found in approximately 5 to 10% of microbial cells in surface marine waters. Strikingly, we also identified a second G1P transporter, UgpB, that is known to transport G3P and belongs to the carbohydrate uptake transporter-1 (CUT1) family, in the model bacterium *Phaeobacter* sp. MED193. To explore the evolutionary pathways that led to the formation of G1P binding sites in both the PhnT and CUT1 families, we determined the structures of GpxB and UgpB bound to G1P and G3P. Using structure-guided mutagenesis and a comparative analysis of the binding pockets within the PhnT and CUT1 families, we traced their evolutionary trajectories, highlighting the distinct strategies through which G1P-binding sites developed in these two protein families.

glycerol-1-phosphate | *Phaeobacter* sp. MED193 | marine bacteria | glycerol-3-phosphate

All cells possess lipid membranes to maintain their integrity and provide essential energy generating mechanisms. Across the three domains of life bacteria and eukaryotes contain membranes comprising fatty acids linked to a sn-glycerol-3-phosphate (G3P) backbone via ester bonds, while Archaea possess isoprenoid hydrocarbon chains linked by ether linkages to a sn-glycerol-1-phosphate (G1P) backbone. Such a divergence in lipid composition has prompted much discussion about the lipid composition of the last universal common ancestor (LUCA) and that LUCA possessed a mixed bacterial/archaeal membrane (1, 2). G1P and G3P represent the core structure of all glycerophospholipids across all domains of life, suggesting these molecules are widely available in all environments.

Bacterial utilization of organic phosphorus sources is a widespread strategy to obtain phosphorus (P) for growth, particularly under conditions when inorganic phosphate is limiting (3). Phospholipids are key constituents of the particulate organic carbon (POC) pool in the global ocean, representing 5 to 30% of POC in surface marine waters (4, 5), yet their utilization and catabolism have received relatively little attention (6, 7). Recently, we provided evidence that bacterial phospholipid headgroups are degraded by marine microbes liberating the nitrogen-containing moieties that are transported across the bacterial membrane and catabolized internally to support microbial growth, highlighting how lipid degradation in the oceans directly links the biogeochemical cycling of P, nitrogen (N), and carbon (C) (6, 7). Surprisingly, the turnover of archaeal lipids in natural environments has been poorly studied, as previous research has primarily focused on their detection as molecular proxies for paleoenvironmental reconstruction in sedimentary settings (8). However, early laboratory microcosm experiments and more recent environmental studies of sediment samples have shown that the ether bonds in archaeal lipids can indeed undergo hydrolysis, leading to the release of glycerol monoether lipids and the corresponding long-chain isoprenoid or nonisoprenoid alcohols (9–11).

While there is a general dearth of data on the diversity and abundance of specific lipid classes in oceanic waters (12, 13), how the Archaea-derived G1P is utilized by marine microbes remains unknown. G1P serves as the fundamental building block for all archaeal membrane lipids, and Archaea play a crucial role in the biogeochemical cycles of C and N in the global ocean, particularly in ammonium oxidation and methane cycling (14, 15).

## Significance

Archaea contribute substantially to global marine biomass, yet how their unique membrane building block, glycerol-1-phosphate, is recycled has remained unexplored. This study identifies two distinct bacterial transport systems, GpxB and UgpB, that mediate high-affinity uptake of both glycerol-1-phosphate and glycerol-3-phosphate. These transporters are widespread and actively expressed in marine microbiomes, indicating that glycerol-1-phosphate cycling represents a previously overlooked link between archaeal biomass and the global carbon and phosphorus cycles. In sharp contrast to the key enzymes involved in glycerol-1-phosphate and glycerol-3-phosphate biosynthesis, these transporter proteins show no chiral selectivity and can bind to both compounds with similar affinity. By uncovering the structural and evolutionary basis of glycerol-1-phosphate transport, this work provides key insights into glycerol phosphate cycling in the oceans.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

<sup>1</sup>N.W., L.M.W., M.L., and C.-Y.L. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Email: zhangyz@sdu.edu.cn, d.j.scanlan@warwick.ac.uk, or y.chen.22@bham.ac.uk.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2524546122/-/DCSupplemental>.

Published December 9, 2025.

Recent estimates suggest that Archaea account for  $10^{28}$  to  $10^{29}$  cells in the global ocean (16, 17) and they alone contribute ~0.3 Gt of marine carbon (18). Given the widespread distribution and abundance of Archaea across Earth's ecosystems (19), we aimed to identify how G1P, the hallmark backbone of archaeal membrane lipids, is used by marine microorganisms.

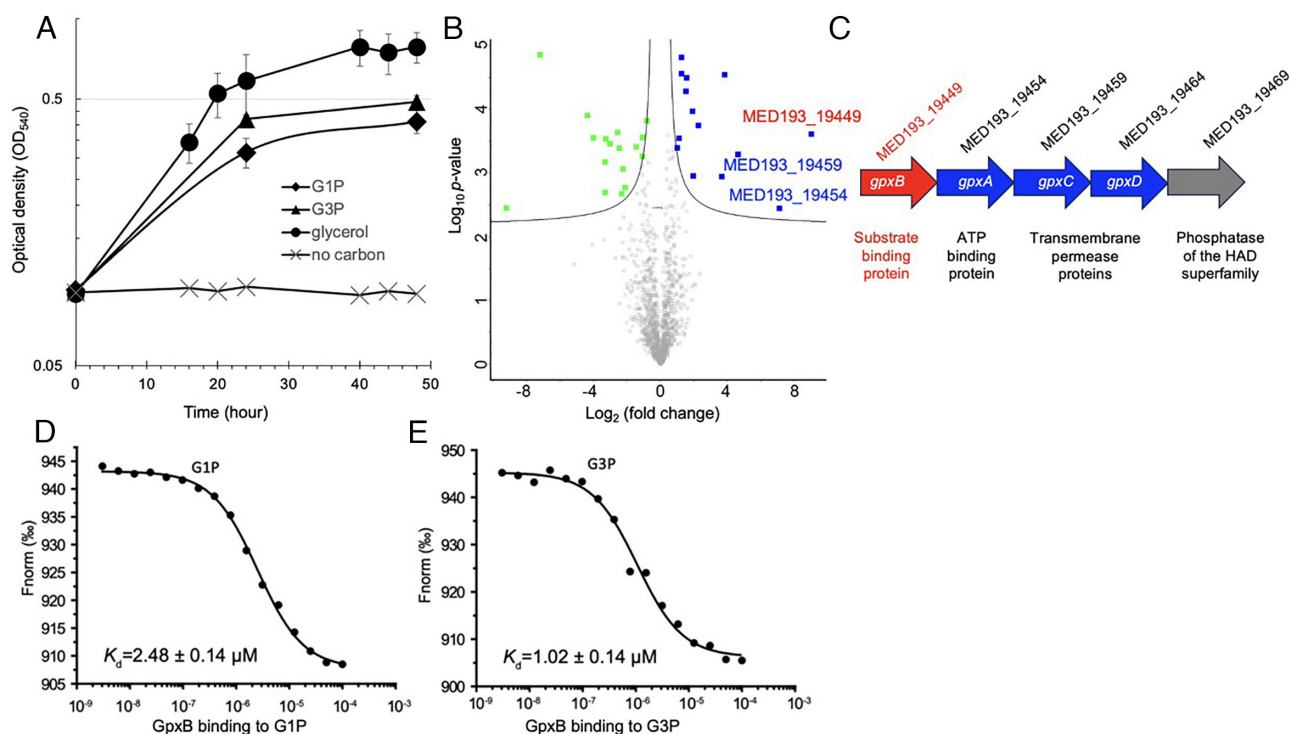
Here, we identify an ABC transporter, termed GpxB, which is a member of the phosphonate uptake transporter family (3.A.1.9) and can transport both G1P and G3P with high affinity. Strikingly, the UgpB transporter, belonging to the carbohydrate uptake transporter family (3.A.1.1), can also transport G1P, despite UgpB and GpxB being distinct protein families with no sequence homology. Through crystallization studies, we elucidated the structural basis for G1P transport by these transporters and propose an evolutionary path explaining GpxB and UgpB selectivity for G1P/G3P in their respective transporter families. Thus, our study not only identifies G1P transporters but also reveals two fundamentally different evolutionary pathways for the development of G1P binding by nature.

## Results and Discussion

**Identification of a G1P/G3P Transporter, GpxB, in Diverse Bacterial and Archaeal Taxa.** In our ongoing efforts to understand lipid metabolism in *Phaeobacter* sp. MED193, a model strain isolated from oligotrophic Mediterranean Sea surface waters (6, 20), we found that this strain grows on G1P as the sole C source (Fig. 1A). It also grows on G3P and glycerol as a carbon source. Thus, *Phaeobacter* sp. MED193 can import and utilize G1P for growth and a G1P transporter must exist in this bacterium. Since a G1P transporter

has not been documented before, we conducted comparative proteomics to uncover the mechanism of import and degradation of the lipid backbone moiety G1P. Thus, *Phaeobacter* sp. MED193 was cultured in G1P, G3P, or glycerol, and the cellular proteomes were compared. The addition of G1P (relative to the addition of glycerol) significantly induced the synthesis of 24 proteins ( $P$ -value  $< 0.05$ ) in the cellular proteome (*SI Appendix, Table S1*). Noticeably, components of a putative phosphonate ABC transporter showed the greatest fold-change (Fig. 1B,  $P < 0.05$ ). The ABC transporter comprised ORFs MED193\_19449 (substrate-binding component, hereafter called GpxB), MED193\_19454 (ATPase component, GpxC) and MED193\_19459 (permease component, GpxD) (Fig. 1C). Interestingly, the same ABC transporter system was also highly upregulated when G3P was supplemented (*SI Appendix, Fig. S1 and Table S1*). Thus, we reasoned this transporter is likely capable of importing both G1P and G3P.

To determine the substrate specificity of the periplasmic binding component encoded by MED193\_19449 (GpxB) we purified recombinant GpxB protein and undertook affinity assays using microscale thermophoresis (MST). Among the array of compounds tested, recombinant GpxB had the highest binding affinity ( $K_d$ ) for G1P ( $K_d = 2.48 \pm 0.14 \mu\text{M}$ ) and G3P ( $K_d = 1.02 \pm 0.14 \mu\text{M}$ ) (Fig. 1D and E). Other 3-carbon ( $\text{C}_3$ ) phosphate substrates tested showed significantly lower or no binding (*SI Appendix, Fig. S2*), suggesting this transporter is specific for G1P/G3P. The strong binding of GpxB to G1P and G3P appears to be a conserved feature, given that three other GpxB proteins from *Sulfitobacter pseudonitzschiae* DSM26824, *Sulfitobacter* sp. EE-36, and *Sulfitobacter geojensis* EhN01 all exhibited low- $\mu\text{M}$  binding affinity toward G1P/G3P (*SI Appendix, Fig. S3*).



**Fig. 1.** GpxB (MED193\_19449) is a G1P binding protein. (A) Growth of *Phaeobacter* sp. MED193 on G1P and G3P. *Phaeobacter* sp. MED193 was cultured in defined artificial seawater medium (ASW) supplemented with 5 mM of either G1P (diamond), G3P (triangle), no carbon control (cross), or glycerol (circle) as the positive control. (B) Volcano plots of cellular proteomics data comparing *Phaeobacter* sp. MED193 grown on G1P and glycerol as the sole C source. GpxB is highlighted in red. Scatter points represent proteins. The x-axis is the fold change for the ratio between both growth conditions and the y-axis is the statistical  $P$ -value. Blue squares represent proteins that are significantly upregulated in the G1P growth condition whereas green squares represent those proteins that are significantly downregulated ( $P$ -value  $< 0.05$ , log<sub>2</sub> fold-change). A complete list of differentially regulated proteins by G1P is shown in *SI Appendix, Table S1*. (C) Gene neighborhood of the *Phaeobacter* sp. MED193 G1P transporter, including the substrate-binding protein GpxB, the ATP binding protein GpxA and the transmembrane proteins GpxC and GpxD. (D and E) Detection of G1P (D) and G3P (E) binding to purified *Phaeobacter* sp. MED193 GpxB by microscale thermophoresis (MST). The change in MST signals was fitted (black lines) to yield a  $K_d$  of 2.48 and 1.02  $\mu\text{M}$  for G1P and G3P respectively.

Genome analysis suggests this G1P transporter appears widely distributed in a variety of bacterial and archaeal taxa, noticeably alphaproteobacterial Rhodobacterales, Hyphomicrobiales, Rhodospirillales, gammaproteobacterial Oceanospirillales, Pseudomonadales, Vibrionales, betaproteobacterial Burkholderiales, and the heterotrophic archaeon Haloferacaceae (SI Appendix, Fig. S4 and Table S2). Analyses of *gpxB* abundance and transcription in the global Tara Ocean database confirmed that it is widely distributed across various Alphaproteobacteria, actively expressed in global ocean microbiomes, and present in approximately 1 out of every 10 to 20 cells in surface ocean environments (SI Appendix, Fig. S5).

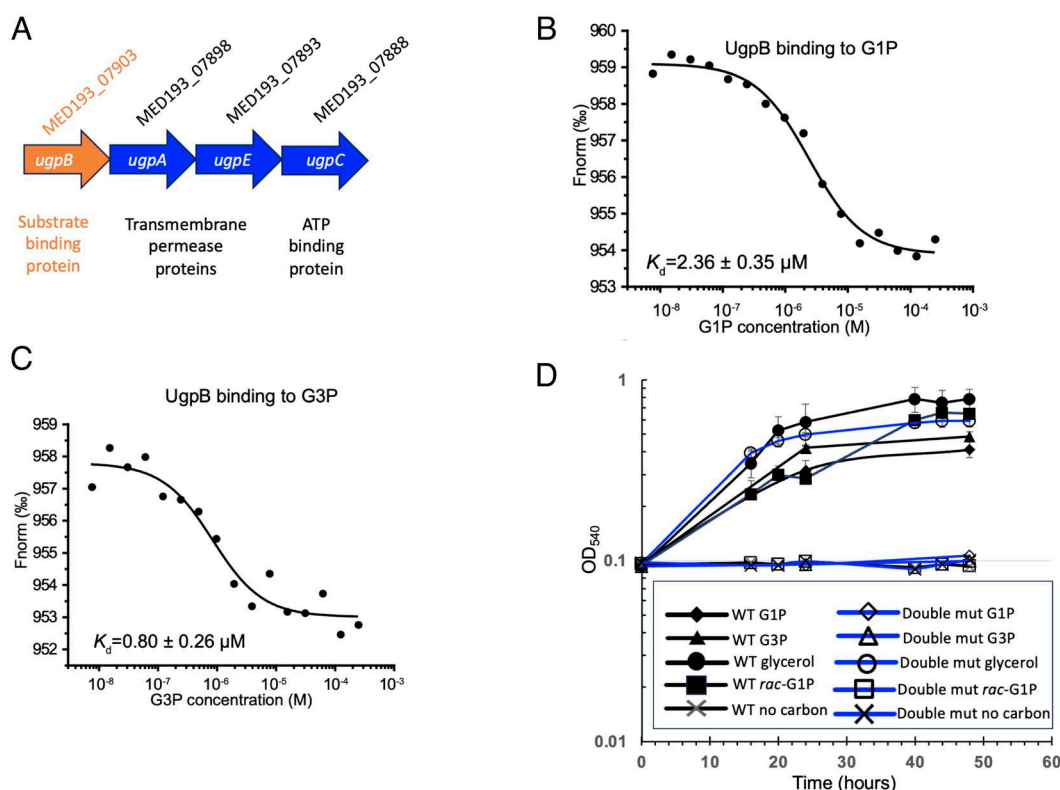
**The Bacterial UgpB Transporter Also Binds G1P.** Given that GpxB binds both G1P and G3P, we next investigated whether the bacterial G3P transporter UgpB (Fig. 2A) is also capable of binding G1P although UgpB has thus far only been shown to transport G3P (21). To assess whether UgpB can bind G1P, we heterologously overexpressed the *Phaeobacter* UgpB homologue (MED193\_07903) and performed binding assays (SI Appendix, Fig. S6). Among the substrates tested, recombinant UgpB had binding affinities ( $K_d$ ) of  $2.36 \pm 0.35 \mu\text{M}$  for G1P and  $0.80 \pm 0.26 \mu\text{M}$  for G3P, suggesting UgpB can transport both G1P and G3P in this bacterium (Fig. 2B and C). GpxB and UgpB appear to be the sole substrate-binding proteins for G1P and G3P in this bacterium, as a  $\Delta\text{ugpB}\Delta\text{gpxB}$  double mutant was unable to grow on either substrate as the sole carbon source, although growth on glycerol remained unaffected (Fig. 2D). These findings collectively indicate that UgpB and GpxB function as the primary transporters for G1P and G3P in this bacterium.

It is striking to note that both glycerol phosphate-binding proteins, UgpB and GpxB, show weak-to-no selectivity against G1P and G3P, with their  $K_d$  values for G1P and G3P differing by less than three-fold. This is in sharp contrast to the G1P and G3P biosynthesis enzymes, G1P dehydrogenase (E.C. 1.1.1.261), and G3P dehydrogenase (E.C. 1.1.1.94), which catalyze the stereospecific reduction of dihydroxyacetone phosphate to form G1P and G3P, respectively (22). Structural and biochemical studies have demonstrated that such chiral selectivity in these glycerol phosphate dehydrogenases arises from the stereochemical mechanism of hydride ion transfer between the substrate (dihydroxyacetone phosphate) and the nicotinamide ring of the reductant NADH. While G3P dehydrogenase transfers the hydride ion from the *si* face of the nicotinamide ring to the substrate, G1P dehydrogenase transfers the hydride ion from the *re* face (23–25).

#### Elucidation of the Contrasting G1P Binding Sites in GpxB and UgpB.

Although both GpxB and UgpB can function as G1P transporters in this bacterium, they lack sequence similarity. Thus, questions arise as to how these proteins may have evolved to bind glycerol phosphates. Substrate-binding proteins are part of a structurally related class, which can be grouped into seven well-defined clades (26, 27). GpxB is classified under the phosphonate uptake transporter (PhnT) family (3.A.1.9) within the D-III clade, with a predicted molecular weight (MW) of 35.0 kDa. In contrast, UgpB belongs to the carbohydrate uptake (CUT1) transporter family (3.A.1.1) of the D-I clade and has a predicted MW of 47.2 kDa.

Since GpxB and UgpB are the only known transporters for G1P to date, to uncover the structural basis of G1P binding in these proteins, we solved both structures using GpxB from the



**Fig. 2.** UgpB is the second G1P binding protein in *Phaeobacter* sp. MED193. (A) Gene neighborhood of the *Phaeobacter* sp. MED193 UgpB transporter, including the substrate-binding protein UgpB, the ATP binding protein UgpC and the transmembrane proteins UgpA and UgpE. (B) *Phaeobacter* sp. MED193 UgpB binding to G1P. The change in MST signals was fitted (black lines) to yield a  $K_d$  of  $2.36 \mu\text{M}$  for G1P. (C) *Phaeobacter* sp. MED193 UgpB binding to G3P. The change in MST signals was fitted (black lines) to yield a  $K_d$  of  $0.80 \mu\text{M}$  for G3P. (D) Growth of the *gpxB* $\Delta$ *ugpB* double mutant (blue lines) and wild type (black lines) *Phaeobacter* sp. MED193 on 5 mM G1P (diamonds) and G3P (triangles) as the sole C source. Glycerol (circles) is used as a positive control and a no carbon negative control (crosses) is also included. *Rac*-G1P is a 1:1 racemic mixture of L-glycerol 3-phosphate (G3P) and L-glycerol-1-phosphate (G1P).

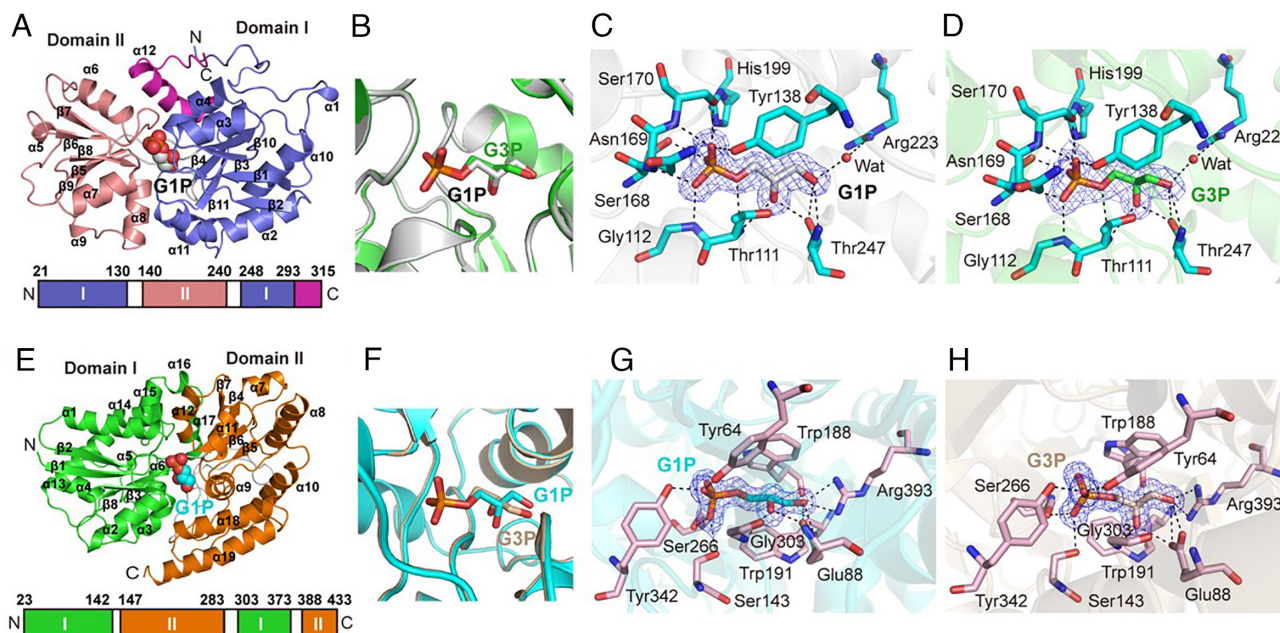


gammaproteobacterium *Marinobacter* sp. DSM11874 (hereafter GpxB<sup>DSM11874</sup>) in complex with G1P and G3P. GpxB<sup>DSM11874</sup> possesses 50% sequence identity to that of *Phaeobacter* sp. MED193 and also showed highest binding toward G1P ( $K_d$   $1.72 \pm 0.28$   $\mu$ M) and G3P ( $K_d$   $1.06 \pm 0.15$   $\mu$ M) among the C<sub>3</sub> phosphates tested (SI Appendix, Fig. S2). Its structure was solved at a resolution of 1.68 Å and 2 Å in complex with G1P and G3P, respectively, with a rmsd of 0.09 Å (SI Appendix, Table S3). The overall structure of GpxB<sup>DSM11874</sup> exhibits the typical topology of type II periplasmic binding proteins (PBPs) (28), with two domains separated by a hinge region containing two loops (residues 131–139 and 241–247) (Fig. 3A). G1P and G3P are aligned in the binding pockets in a similar manner (Fig. 3B) with the phosphate “head” group coordinated by Gly112, Thr111, Tyr138, Ser168, Asn169, Ser170, and His199, whereas the two hydroxyl groups on the “tail” end of G1P and G3P are coordinated through hydrogen bonding with Thr111, and Thr247 (Fig. 3C and D). Additionally, the side chain of Arg223 forms another hydrogen bond with the C3 hydroxyl group of G1P or C1 hydroxyl group of G3P, respectively, via a water molecule (Fig. 3C and D). Site-directed mutagenesis experiments confirmed the crucial role of both phosphate-coordinating residues (e.g., Tyr138, His199) and hydroxy groups-coordinating residues (Thr247, Arg223) in G1P binding and the corresponding alanine mutants exhibited a 10- to 100-fold reduction in binding affinity (SI Appendix, Fig. S7).

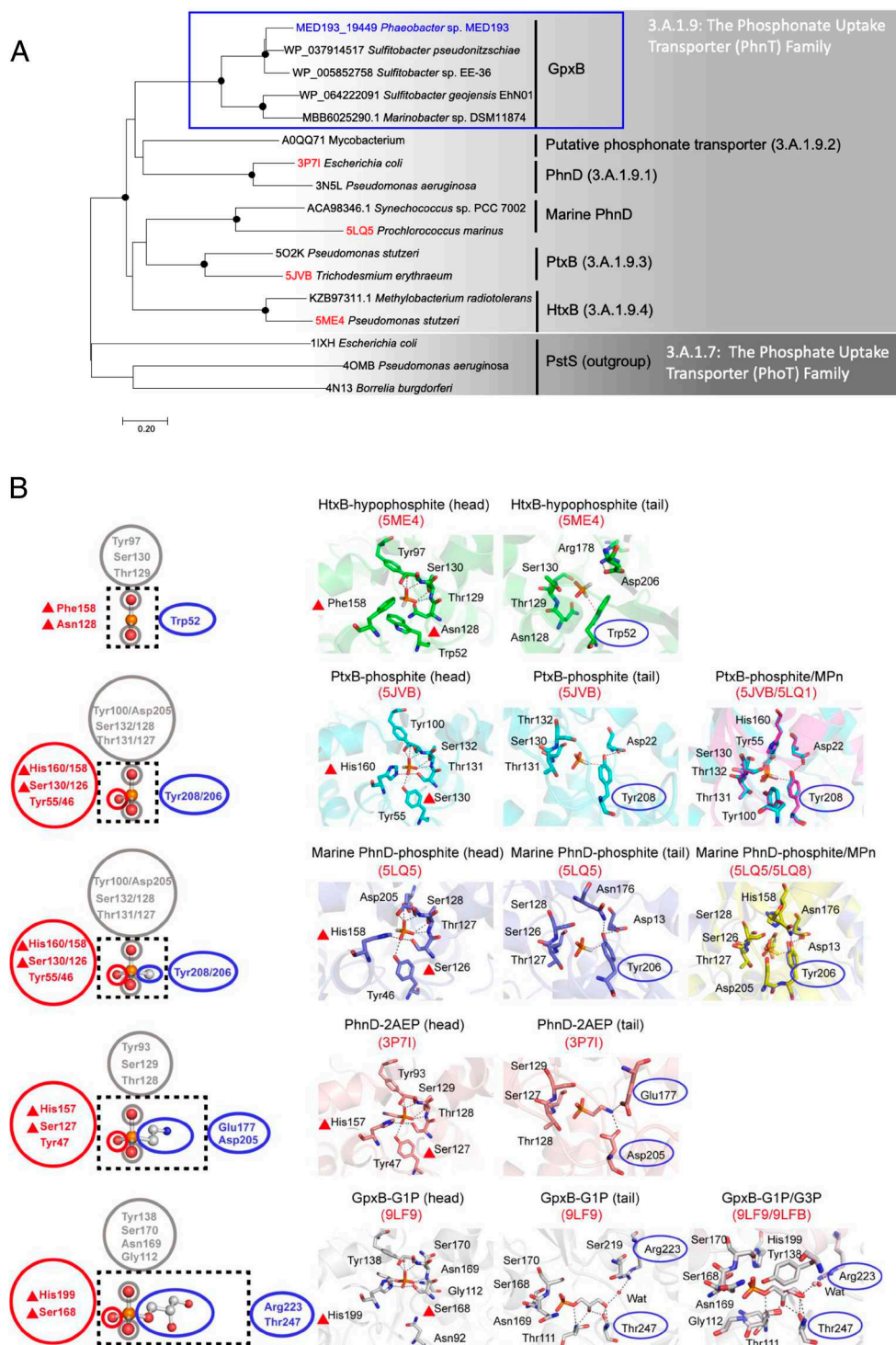
The UgpB structure of *Phaeobacter* sp. MED193 was also solved in the presence of G1P and G3P ligands (SI Appendix, Table S3). The overall structure of UgpB contains two  $\alpha/\beta$  domains connected

via three hinges (residues 143–146, 284–302, and 374–387) and a G1P molecule occupies the active-site between the two domains. Domain I (residues 23–142 and 303–373) and domain II (residues 147–283 and 388–433) are composed of a four-stranded  $\beta$ -sheet surrounded by 11  $\alpha$ -helices and a four-stranded  $\beta$ -sheet surrounded by 8  $\alpha$ -helices, respectively (Fig. 3E). Again, both G1P and G3P occupy the same binding pockets in UgpB (Fig. 3F). The two hydroxyl groups of G1P/G3P are in close contact with Gly303, Arg393 and Glu88 by hydrogen bonds whereas the phosphate head group is coordinated by Tyr64, Ser143, Ser266, and Tyr342 (Fig. 3G and H). A noticeable feature of the UgpB binding pocket, that makes it distinct from GpxB, is the presence of two highly conserved Trp residues (Trp188, Trp191) in UgpB which provides  $\pi$ -stacking interactions and forms a clamp holding the glycerol moiety of the linear G1P/G3P (Fig. 3G and H) (21). Moreover, Trp191 also forms O-H... $\pi$  hydrogen bonding with the oxygen atom of the C2 hydroxyl group of G3P (3.4 Å) (29) (Fig. 3H). Indeed, mutagenesis analyses confirmed the key role of all these Trp residues in securing G1P and G3P in its binding pocket (SI Appendix, Fig. S8). Thus, although both UgpB and GpxB can bind to G1P (and G3P), their binding pockets are strikingly different.

**Emergence of G1P-Binding Sites in the PhnT Family.** Identification of these two contrasting G1P-binding proteins in two different families of transporters thus provides a unique opportunity to uncover the trajectory of the formation of two contrasting G1P binding sites from a phylogenetic and evolutionary perspective. Phylogenetic analyses placed GpxB on a distinct branch within the PhnT family (Fig. 4A), together with substrate-binding proteins



**Fig. 3.** Structural comparison of GpxB<sup>DSM11874</sup> and UgpB binding to G1P. (A) The overall structure of the GpxB<sup>DSM11874</sup>-G1P complex (Top). Domain I, domain II, and the C-terminal helix are colored in slate, salmon, and magenta, respectively. The G1P ligand in GpxB<sup>DSM11874</sup> is shown as gray (carbon), red (oxygen), and orange (phosphorus) spheres. Block diagram representation (Bottom) of each domain in GpxB<sup>DSM11874</sup> colored as in the top. The amino acids that determine the border of each domain are labeled with numbers. (B) Structural overlay of the binding pocket of the GpxB<sup>DSM11874</sup>-G1P complex and GpxB<sup>DSM11874</sup>-G3P complex. G3P is shown as green (carbon), red (oxygen), and orange (phosphorus). (C) A detailed view of the GpxB<sup>DSM11874</sup>-G1P binding sites. G1P is colored in gray and GpxB<sup>DSM11874</sup> residues involved in binding G1P are colored in cyan. The water molecule is displayed as a red sphere. The 2F<sub>o</sub> - F<sub>c</sub> densities for the G1P molecule are contoured in blue at 1.0σ. Dashed lines represent possible hydrogen bonds. (D) A detailed view of GpxB<sup>DSM11874</sup> G3P binding sites. G3P is colored in green and GpxB<sup>DSM11874</sup> residues involved in binding G3P are colored in cyan. The water molecule is displayed as a red sphere. The 2F<sub>o</sub> - F<sub>c</sub> densities for G3P are contoured in blue at 1.0σ. Dashed lines represent possible hydrogen bonds. (E) The overall structure of the UgpB-G1P complex (Top) highlighting domain I (green) and domain II (orange) with the G1P molecule in UgpB. Block diagram representation (Bottom) of each domain in UgpB colored as in the top. The amino acids that determine the border of each domain are labeled with numbers. (F) Structural superposition of the binding pocket of the UgpB-G1P complex (cyan) and the UgpB-G3P complex (wheat) with the G1P and G3P ligands shown as sticks. (G) Residues involved in binding G1P in UgpB. The G1P molecule is shown as cyan sticks and UgpB residues as pink sticks. The 2F<sub>o</sub> - F<sub>c</sub> densities for the G1P molecule are contoured in blue at 1.0σ. Possible hydrogen bonds are drawn as dashed lines. (H) Residues involved in binding G3P in UgpB. The G3P molecule is shown as wheat sticks and UgpB residues as pink sticks. The 2F<sub>o</sub> - F<sub>c</sub> densities for G3P are contoured in blue at 1.0σ. Possible hydrogen bonds are drawn as dashed lines.



**Fig. 4.** Evolution of glycerol phosphate binding in the PhnT family. (A) Phylogenetic tree showing the relationship between GpXB proteins of the PhnT family (3.A.1.9). In the Transporter Classification Database (<https://www.tcd.org/>), four members of the PhnT family associated with organic phosphonate transport are known, including C-P lyase linked PhnD (3.A.1.9.1), an uncharacterized *Mycobacterium* phosphonate transporter (3.A.1.9.2), a phosphite transporter (PtxB, 3.A.1.9.3), hypophosphite transporter (HtxB, 3.A.1.9.4), as well as unclassified PhnD homologues in marine *Synechococcus* and *Prochlorococcus* (marine PhnD). PstS phosphate binding proteins belonging to the phosphate uptake transporter family (3.A.1.7) are used as the outgroup. (B) Schematic plot (Left) and detailed view (Right) of the evolution of phosphonate binding pockets and key residues coordinating the phosphorus-containing head and the non-phosphorus-containing tail. Key evolving residues involved in binding the head and tail of the phosphonate are highlighted with red triangles and blue ovals, respectively. As the head evolves from hypophosphite to phosphite groups, a conserved histidine and a conserved serine (red triangles) evolved to help stabilize the oxygen atom of the third P-O binding site. Meanwhile, an increase in volume of the binding pockets appears while the tail grows from a single H atom to more complex hydroxy groups in G1P and G3P (SI Appendix, Table S4). As the tail evolves from a single H atom to a -CH<sub>3</sub> group, a conserved aromatic amino acid residue evolved for hydrogen bonding with the hydrogen atom in (hypo)phosphite or methylphosphonate through a P-H... $\pi$  or a C-H... $\pi$  bond (Trp52 in HtxB, Tyr208 in PtxB, Tyr206 in marine PhnD). Finally, a pair of Arg223/Thr247 residues in GpXB evolved to orient the two hydroxyl groups in G1P/G3P, while a pair of conserved Glu177/Asp205 residues in PhnD evolved to coordinate the positively charged -NH<sub>3</sub><sup>+</sup> group in 2AEP. In the schematic plot, the ligands are displayed as ball-and-stick and black dashed boxes represent the substrate-binding pockets. Residues in red and gray circles are involved in binding the phosphorus-containing head. In the detailed view of the ligand binding sites, proteins are shown as cartoons, with the bound ligands and selected residues shown as sticks. The water molecule is displayed as a red sphere. Black dashed lines represent possible hydrogen bonds and P-H... $\pi$  and C-H... $\pi$  interactions are drawn as red dashed lines.

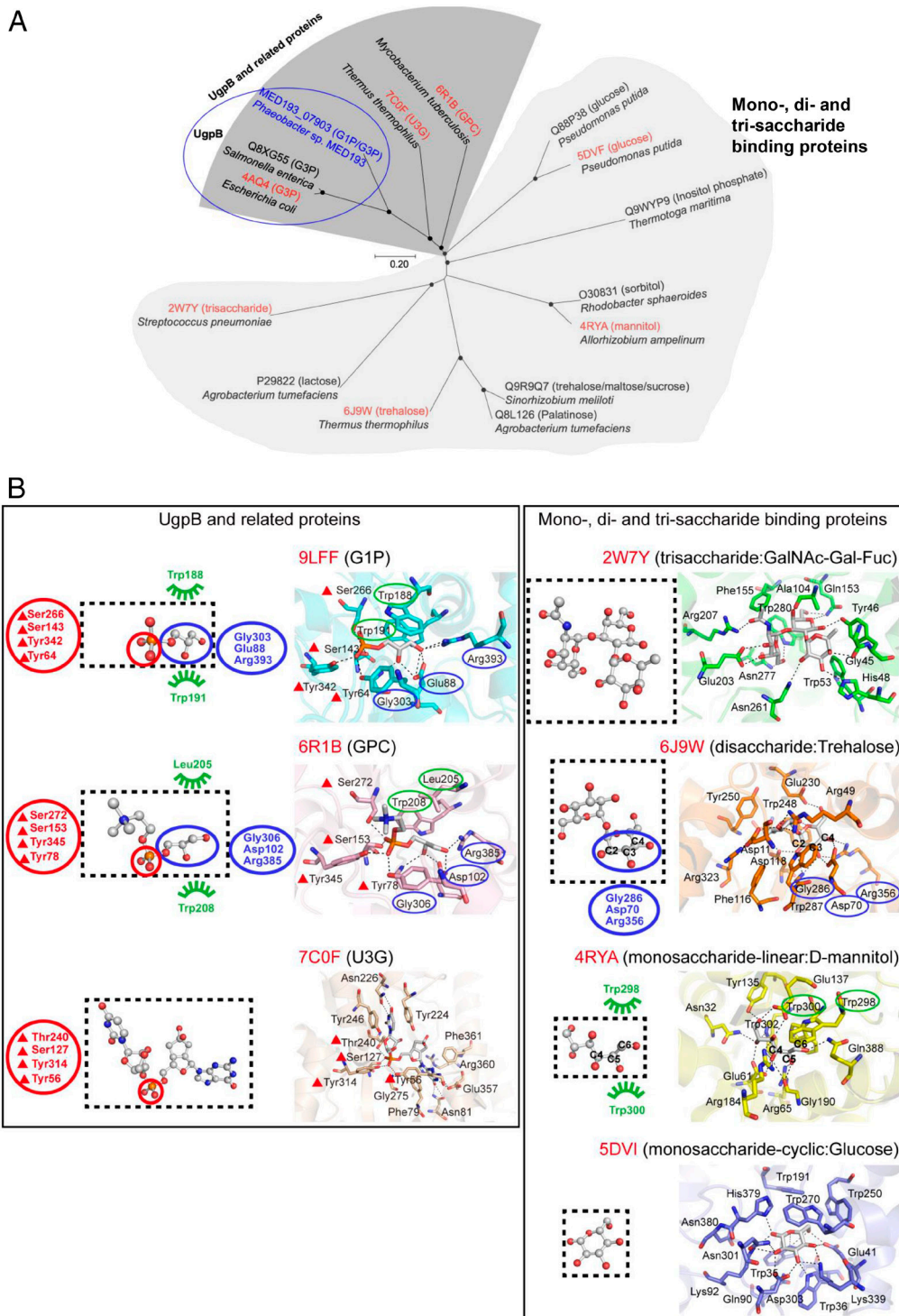


involved in transporting hypophosphite ( $\text{H}_2\text{PO}_2\text{H}$ ), phosphite ( $\text{H}_2\text{PO}_3\text{H}$ ), and small phosphonates like methylphosphonate ( $\text{CH}_3\text{-PO}_3\text{H}_2$ ), 2'-aminoethylphosphonate (2AEP,  $\text{NH}_2\text{-CH}_2\text{-CH}_2\text{-PO}_3\text{H}_2$ ). While the P-containing head of these substrates remains similar in size and charge, the tail group varies significantly, expanding from single H atoms in (hypo)phosphite to larger, more complex aminoethyl- and dihydroxypropanone groups in phosphonates. Correspondingly, the binding pockets of these proteins have evolved to accommodate this increase in tail size. Indeed, GpxB, which binds to G1P/G3P, has the largest binding pocket, while HtxB, which binds hypophosphite, has the smallest (Fig. 4B and *SI Appendix, Table S4*). An evolutionary pathway for key residues involved in coordinating head and tail becomes evident through comparative structural analysis of PhnT family proteins. For example, while residues involved in binding the P-containing head are largely conserved, a new histidine and a serine emerged in PtxB, PhnD, and GpxB (red triangles, Fig. 4B and *SI Appendix, Fig. S9A*) to coordinate the third oxygen moiety as the head evolves from two P-O binding sites (hypophosphite) to three P-O binding sites (phosphite, G1P/G3P, and other phosphonates). Similarly, in HtxB and PtxB, which binds hypophosphite and phosphite, respectively (30), a crucial P-H $\cdots\pi$  hydrogen bond orients the hydrogen atom attached to P in (hypo)phosphite, with a conserved aromatic residue playing a key role (Trp52 in HtxB and Tyr208 in PtxB). This aromatic tyrosine cap appears to help stabilize the closed state of the protein and acts as a steric barrier for blocking the binding of ligands with larger R groups (30). However, in GpxB, such aromatic residues do not exist in corresponding positions (*SI Appendix, Fig. S9A*). As the tail of the substrate expands from a hydrogen atom to a positively charged amine (in 2AEP) or hydroxyl groups (in G1P and G3P), an increase in volume of the binding pockets appears along with the loss of the steric barrier tyrosine. Two key residues in PhnD, Glu177/Asp205, evolved to accommodate the  $\text{-NH}_2$  group (31), whereas in GpxB, these residues have been replaced by Arg223/Thr247, which form hydrogen bonds with the hydroxy groups in G1P and G3P (Fig. 4B and *SI Appendix, Fig. S9A*). This suggests that GpxB likely evolved from smaller (hypo)phosphite-binding proteins, with gradual changes in key residues to accommodate the increasing size, charge, and polarity of the tail group, ultimately enabling specific binding to G1P and G3P.

**Emergence of G1P-Binding Sites in the CUT1 Family.** The CUT1 family of transporters can bind a variety of tri-, di-, and monosaccharides, as well as phosphorylated glycerols and their derivatives. Unsurprisingly, the evolution of the CUT1 family appears more complex. One notable evolutionary pathway for some members involves the progressive shrinking of binding pockets, transitioning from trisaccharides to disaccharides and then to monosaccharides, as shown in Fig. 5A and *SI Appendix, Table S5*. The formation of monohexose-binding pockets likely marks a key branching point in the evolution of the CUT1 family, given the existence of both open-chain and cyclic forms of monohexose. A comparison of the ligands of these proteins shows some similarities between the glycerol moiety (two adjacent hydroxyl groups) of G1P/G3P and the hydroxy groups of the first sugar rings in trehalose and those of D-mannitol (Fig. 5A). There are two likely evolutionary scenarios from sugar-binding proteins to glycerophosphate-binding proteins. In the first scenario, UgpB may evolve from oligosaccharide-binding proteins by retaining residues engaged in hydrogen bonding from the hydroxyl oxygen (Fig. 5B). Thus, in agreement with previous speculation (32), key residues involved in orienting the two hydroxy groups in UgpB (Gly303, Glu88 and Arg393) are conserved in the trehalose-binding protein (Gly286, Asp70 and Arg356). In support of this

hypothesis, UgpB and the hexitol-binding protein (D-mannitol binding protein 4RYA) both contain two conserved Trp residues (188/191 in UgpB and 298/300 in the hexitol-binding protein) involved in forming a clamp to hold the glycerol moiety of the linear substrate G1P/G3P and D-mannitol (Fig. 5B). In the second scenario, UgpB may evolve from an open-chain monohexose-binding protein by incorporating phosphate as a “new” functional group in the substrate. Phylogenetically, UgpB, along with several other members of the CUT1 family, forms a distinct branch separate from proteins that bind disaccharides and oligosaccharides (Fig. 5B). A closer examination of the multiple sequence alignment reveals two proteins closely related to UgpB: one from *Mycobacterium tuberculosis*, the glycerophosphocholine (GPC)-binding protein (33), and the other the uridylyl-3'-5'-phosphoguanosine (U3G)-binding protein (32), both of which contain a phosphate group (Fig. 5B). Remarkably, two Tyr residues and two Ser residues that coordinate the phosphate P-O bond (red triangles) are strictly conserved in UgpB, as well as in the GPC and U3G-binding proteins (*SI Appendix, Fig. S9B*), but not in glucose-binding proteins. Moreover, the residues forming hydrogen bonds with the hydroxyl oxygen of the glycerol moiety still exist in GPC-binding proteins (Gly306, Asp102 and Arg385) (Fig. 5B). During the evolution from GPC and U3G-binding proteins to UgpB, the volume of their binding pockets shrinks to accommodate the decrease in ligand size (*SI Appendix, Table S5*), and at the same time, two conserved Trp residues (188/191) appear to form a clamp to hold the glycerol moiety of the linear G1P/G3P (Fig. 5B). This evolution results in the reshaping of the binding pocket to accommodate the negatively charged phosphate moiety and supports the hypothesis that UgpB may have evolved from a non-phosphate-containing linear hexose-binding protein.

Interestingly, structural comparison of the GpxB and UgpB proteins also provides insights into the evolutionary trajectories of the PhnT and CUT1 protein families. Proteins in the PhnT family appear to evolve from a phosphorus-containing head toward a non-phosphorus-containing tail, whereas the CUT1 family exhibits the opposite trend, evolving from non-phosphorus-containing structures toward phosphorus-containing heads. In the binding cavity of GpxB, five water molecules are located near the tail of G1P (*SI Appendix, Fig. S10A*). Similarly, five water molecules were observed behind the  $\text{-NH}_2$  group of 2AEP in the binding cavity of *EcPhnD* (31). These water molecules may create a spatial foundation for binding phosphonates/organophosphorus compounds with longer R1 groups during the evolution of the PhnT family (*SI Appendix, Fig. S10A*). Conversely, in UgpB, five water molecules occupy the head region of the phosphorus-containing G1P within the binding cavity, potentially supporting an evolutionary transition from a glycerol sidechain to a phosphorus-containing head, and ultimately to the choline moiety of GPC (*SI Appendix, Fig. S10B*). The binding affinity of GpxB and UgpB is further reflected in their hydrogen bond networks. GpxB exhibits a more extensive network, with 13 hydrogen bonds stabilizing the binding of G1P, including seven hydrogen bonds specifically with the phosphate head group (Fig. 3C). In contrast, UgpB forms 10 hydrogen bonds in total, with only four engaging the phosphate head group (Fig. 3G). The denser hydrogen bond network in GpxB suggests it likely binds the phosphate group more strongly than UgpB. Additionally, residues in UgpB's binding pocket play a particularly significant role at the individual residue level, as demonstrated by site-directed mutagenesis experiments (*SI Appendix, Figs. S7 and S8*). This highlights the structural importance of these residues for UgpB's binding function. Thus, while both GpxB and UgpB can bind to G1P (and G3P), they exhibit striking structural and functional differences that reflect their distinct evolutionary and biochemical characteristics.



**Fig. 5.** Evolution of glycerol phosphate binding in the CUT1 family. (A) A phylogenetic tree illustrating the relationships between UgpB (highlighted in a blue oval) and closely related members (highlighted in shaded gray) of the Carbohydrate Uptake Transporter-1 (CUT1) Family (3.A.1.1). (B, Left) A close examination of UgpB and its related members in the CUT1 family reveals connections to a GPC binding protein from *M. tuberculosis* (33) and a U3G binding protein from *Thermus thermophilus* (32). There is a gradual increase in the size of the binding pockets from G1P/G3P binding proteins to GPC and U3G binding proteins (SI Appendix, Table S5). The evolution of key residues from the phosphate-containing U3G to GPC to G1P/G3P is also illustrated, highlighting two conserved Tyr residues (64/342) and two Ser residues (143/266) involved in orienting the oxygen atoms of P-O in phosphate (red triangles). In addition, two conserved Trp residues (188/191) emerged in binding the glycerol moiety of the linear G1P/G3P in UgpB (green ovals). (Right) The evolution of binding pockets depicted by surface representation of proteins demonstrating a transition from trisaccharides to disaccharides to monohexose, along with a decrease in the volume of the substrate-binding pockets (SI Appendix, Table S5). Noticeably, the two hydroxyl groups in G1P/G3P are akin to the hydroxy groups of C2-C4 of the first glycosyl unit of trehalose as well as those in D-mannitol. A comparison of the binding pockets of trehalose-binding protein and G1P/G3P binding UgpB shows that three residues (Gly286, Asp70, Arg356, and Gly303, Glu88, Arg393 respectively) involved in coordinating the oxygen atoms of the glycerol moiety are highly conserved (blue ovals). Similarly, the binding pocket of UgpB also contains two conserved Trp residues (188/191 in UgpB and 298/300 in the D-mannitol-binding protein 4RYA) involved in binding the glycerol moiety of the linear G1P/G3P (green ovals). In the schematic plot, ligands are displayed as ball-and-stick and black dashed boxes represent the substrate-binding pockets. Residues in red and gray circles are involved in binding the phosphorus-containing head. In the detailed view of the ligand binding sites, proteins are shown as cartoons with the bound ligands and selected residues shown as sticks. Key evolving residues involved in binding the phosphate group and glycerol moiety are highlighted with triangles and ovals. Possible hydrogen bonds are drawn as dashed lines.



## Conclusions

In summary, we have identified two G1P transporters and elucidated the molecular mechanisms underlying their high-affinity binding to G1P. Our findings highlight a remarkable instance of functional convergent evolution, showcasing how two seemingly unrelated proteins have adapted to bind the same substrate to support microbial growth. High-affinity transporters for G1P are crucial for its uptake, and further investigation into the regulatory mechanisms governing GpxB and UgpB will provide deeper insights into the roles of these transporters in G1P metabolism. Given the vast abundance of Archaea in the global oceans, G1P cycling is likely to be particularly pronounced in these waters. The prevalence of G1P transporter proteins in the global ocean microbiome suggests that G1P cycling plays a significant yet often overlooked role in the global C and P cycles.

## Materials and Methods

**Strains and Culture Conditions.** *Phaeobacter* sp. MED193 and its mutant strains were routinely maintained on 1.5% (w/v) Difco marine broth 2216 agar with antibiotics as appropriate. To test the growth of *Phaeobacter* sp. MED193 on G1P/G3P, cells were first cultivated in 5 mL marine broth medium overnight at 30 °C with shaking at 150 rpm. A 1% (v/v) inoculum was then transferred to defined artificial seawater medium supplemented with 50  $\mu$ M phosphate (6). After 24 h growth, 12.5 mL cultures were collected by centrifugation and inoculated into 100 mL artificial seawater medium. Subsequently, either G1P or G3P (purchased from Sigma-Aldrich, St. Louis, USA) was added to a final concentration of 5 mM as the sole C source. Cells (40 mL) were harvested after 48 h and used for proteomics.

All *Escherichia coli* strains were routinely cultivated in LB medium supplemented with antibiotics as appropriate. To express G1P/G3P substrate-binding proteins, genes were chemically synthesized (see below) and subcloned into pET-22b or pET-28a vectors and transformed into *E. coli* BL21 (DE3). *E. coli* cells were routinely cultivated in LB medium at 37 °C until the OD<sub>540</sub> reached 0.3 to 0.5 before protein expression was induced by adding 0.1 to 0.5 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) (final concentration) and cultures shifting to 15 to 22 °C for a further 6 to 12 h before harvesting by centrifugation.

**Proteomics Analysis.** For cellular proteomes, cell pellets were resuspended in 1 mL ddH<sub>2</sub>O and 200  $\mu$ L LDS buffer was added. For exoproteomes, proteins were precipitated from cell-free culture medium using trichloroacetic acid as described previously (34). Cellular and exo protein samples were then boiled at 98 °C for 15 min before loading onto a precast NuPAGE Bis-Tris gel for a short run at 140 V for 5 to 15 min. After staining with Coomassie blue and destaining overnight in ddH<sub>2</sub>O, protein bands were cut into small pieces and extracted as described previously (34). Trypsin digested peptides were cleaned using Costar Spin-X centrifuge tube filters and analyzed by nanoLC (Ultimate 3000)-ESI MS/MS (Orbitrap Fusion mass spectrometer, ThermoFisher) at the University of Warwick Research Technology Platform. The recorded MS/MS data was processed using MaxQuant (v1.5.5.1) and Perseus (v1.6.5.0). Statistical analysis was set using a false discovery rate of 0.05 and a minimal log<sub>2</sub> fold change of two. Only proteins that were present in every biological replicate of at least one condition (G1P, G3P, or glycerol) were retained for statistical analyses.

**Microscale Thermophoresis.** The binding affinity of the purified proteins to G1P and other structurally related chemicals was measured using a Monolith NT.115 (NanoTemper Technologies, Munich, Germany). Proteins (160 nM) were labeled in assay buffer (PBS containing 0.05% (v/v) Tween-20) using the Protein Labelling Kit RED-Tris-NTA (NanoTemper Technologies). For binding assays, the buffer was replaced by 20 mM HEPES pH 7.4, 100 mM NaCl, 0.05% (v/v) Tween-20. Each chemical was prepared by serial dilution in the same assay buffer and mixed with the same volume of labeled protein (at a final concentration of 80 nM). The mixture was then loaded onto Monolith™ NT.115 Series capillaries (NanoTemper Technologies, Germany) and thermophoresis was measured at 22 °C using 60% excitation power and medium MST power. Data were analyzed using MO. Affinity Analysis v2.3 software (NanoTemper Technologies).

**Construction of *Phaeobacter* Mutants.** The *gpxB**ugpB* double mutant was generated using a conjugation protocol as described previously (35). First, two homologous regions of an upstream and a downstream fragment flanking the *gpxB* gene were cloned into suicide plasmid pK18*mobsacB*, together with a spectinomycin (Spec) resistance cassette which was inserted in between the upstream and downstream fragments. The resulting plasmid was then mobilized into *Phaeobacter* sp. MED193 using *E. coli* S17.1 lambda *pir* and transconjugants selected on marine broth MB agar containing Spec (200  $\mu$ g mL<sup>-1</sup>). Putative mutants were confirmed by replicate plating on MB agar with Spec and Kan respectively and those colonies that did not grow on Kan plates were further validated by PCR using primers targeting the gene of interest. The same protocol was repeated to generate the *gpxB**ugpB* double mutant in the *gpxB* single mutant background, except that a Gentamycin (Gm) resistant cassette was inserted in between the upstream and downstream fragments of *ugpB* into pK18*mobsacB*. All primers, plasmids, and bacterial strains used are listed in *SI Appendix, Table S6*.

**Gene Synthesis, Point Mutations, and Protein Overexpression and Purification.** The full-length codon optimized GpxB (MED193\_19449) of *Phaeobacter* sp. MED193 (WP\_009811375.1), and four GpxB homologs (*Marinobacter* sp. DSM11874, MBB6025290.1; *S. pseudonitzschiae* DSM26824, WP\_037914517.1; *Sulfitobacter* sp. EE-36, WP\_005852758.1; *S. geojensis* sp. EhN01, WP\_064222091.1) were chemically synthesized by Beijing Tsingke Biotech Co., Ltd. (China), and subcloned into pET-22b (Novagen, America) with a C-terminal His-tag. An N-terminal signal peptide truncated [residues 1–22 predicted by SignalP 5.0 Server (36)] MED193\_07903 gene encoding UgpB from *Phaeobacter* sp. MED193 (WP\_009809101.1) was synthesized by Beijing Tsingke Biotech Co., Ltd. (China), and then subcloned into pET-28a (Novagen, America) providing an N-terminal His-tag. All point mutations of GpxB<sup>DSM11874</sup> and UgpB were generated via a PCR-based method using PfuUltra II Fusion HS DNA Polymerase (Agilent Technologies, America) and verified by DNA sequencing. All proteins were expressed in *E. coli* BL21 (DE3). *E. coli* cells were routinely cultivated in LB medium supplemented with 100  $\mu$ g mL<sup>-1</sup> ampicillin (30  $\mu$ g mL<sup>-1</sup> kanamycin for UgpB and its mutants) at 37 °C until the OD<sub>600</sub> reached 0.8 to 1.0 before protein expression was induced by adding 0.5 mM IPTG and cultures shifting to 17 °C for a further 14 to 16 h before harvesting by centrifugation. Proteins were first purified by affinity chromatography with Ni<sup>2+</sup>-nitrilotriacetic acid (NTA) resin (Qiagen, Germany) and then fractionated by a Superdex 200 column (Cytiva, America) with buffer containing 10 mM Tris-HCl (pH 8.0), 100 mM NaCl.

**Crystallization and Data Collection.** The purified proteins used for crystallization were prepared to 7 mg mL<sup>-1</sup> (GpxB<sup>DSM11874</sup>) and 8 mg mL<sup>-1</sup> (UgpB) in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl. To obtain the structure of the GpxB<sup>DSM11874</sup>-G1P complex, GpxB<sup>DSM11874</sup> was cocrystallized with 2 mM G1P using the sitting-drop vapor diffusion method at 18 °C for 2 wk using a protein:reservoir solution ratio of 1:1. Crystals of the GpxB<sup>DSM11874</sup>-G1P complex were obtained in sitting drops containing 0.04 M potassium phosphate, 16% (w/v) polyethylene glycol (PEG) 8000 and 20% (v/v) glycerol. Crystals of the UgpB-G1P complex were obtained using the same cocrystallization method, but in 0.1 M PCB buffer (the molar ratios of sodium propionate to sodium cacodylate to BIS-TRIS propane were 2:1:2) (pH 5.0), 25% (w/v) PEG 1500. Crystals of the UgpB-G3P complex were obtained using the same cocrystallization method, but in conditions containing 0.1 M MIB buffer (the molar ratios of sodium malonate to imidazole to boric acid were 2:3:3) (pH 7.0), 25% (w/v) PEG 1500 in the presence of 2 mM G3P. For the GpxB<sup>DSM11874</sup>-G3P complex, GpxB<sup>DSM11874</sup> was cocrystallized with 2 mM G3P using a hanging-drop vapor diffusion method in crystallization buffer containing 0.1 M calcium acetate, 0.1 M sodium acetate (pH 4.5), 10% (w/v) PEG 4000 at 18 °C for 2 wk.

The X-ray diffraction data of GpxB<sup>DSM11874</sup>-G3P complex was collected by X-ray single crystal diffractometry at the State Key Laboratory of Microbial Technology of Shandong University and initial diffraction datasets were processed using the HKL3000 program with default settings (37). All other X-ray diffraction data were collected on the BL18U1 and BL19U1 beamlines at the Shanghai Synchrotron Radiation Facility.

**Structure Determination and Refinement.** Crystals of the GpxB<sup>DSM11874</sup>-G1P/G3P complex and the UgpB-G1P/G3P complex belonged to the I23 and P2<sub>1</sub>2<sub>1</sub>2 space group. The crystal structure of the GpxB<sup>DSM11874</sup>-G1P complex was determined by molecular replacement using a CCP4 program Phaser (38) with the structure of GpxB<sup>DSM11874</sup> generated by AlphaFold2 (39) as the search model.

The structure of the GpxB<sup>DSM11874</sup>-G3P complex was further determined using the same method, but with the solved GpxB<sup>DSM11874</sup>-G1P structure as a model. The crystal structures of the UgpB-G1P complex and UgpB-G3P complex were solved using the same method. Several rounds of refinement of all structures were carried out in Coot (40) and Phenix (41) using default parameters. All structure figures were processed using the PyMOL program (<http://www.pymol.org/>). The pocket volume of proteins was calculated using PyVOL 1.7.6 (42).

**Circular-Dichroism (CD) Spectroscopic Assays.** GpxB (and its mutants) and UgpB (and its mutants) were prepared to ~0.4 mg mL<sup>-1</sup> in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl. CD spectroscopic assays for these proteins were conducted on a J-1500 Spectrometer (Jasco, Japan) at 25 °C. All the data were collected under the following conditions: wavelength from 250 to 200 nm, scanning speed 500 nm min<sup>-1</sup>, bandwidth 1 nm, data pitch 0.5 nm and data interval 0.5 nm.

**GpxB and UgpB in Bacterial Genomes and the Global Ocean Metagenomes and Metatranscriptomes.** *Phaeobacter* sp. MED193 GpxB was used to search for its homologues in bacterial and archaeal genomes in the NCBI nonredundant protein sequences database using 1e<sup>-50</sup> and sequence coverage >75% as the cut-off value. Tara Oceans metagenomes (OM-RGCv2+G) and metatranscriptomes (OM-RGCv2+T) were queried for GpxB and UgpB via the Ocean Gene Atlas (OGA) web portal as described previously (6). GpxB and UgpB abundance in both metagenomes and metatranscriptomes was obtained by using hmmsearch with an expected threshold of 1e<sup>-80</sup> normalized to the median abundance of 10 single-copy marker genes as described previously (43).

**Data, Materials, and Software Availability.** All structural data generated in this study has been deposited in the Protein Data Bank (PDB) database ([www.pdb.org](http://www.pdb.org)) with accession code 9LF9 (44), 9LFB (45), 9LFF (46), and 9LEJ (47) for GpxB<sup>DSM11874</sup>-G1P, GpxB<sup>DSM11874</sup>-G3P, UgpB-G1P, and UgpB-G3P, respectively.

**ACKNOWLEDGMENTS.** This work was supported by the Natural Environment Research Council (NERC, UK) via a PhD scholarship (The Central England NERC Training Alliance 2, CENTA2, NE/S007350/1) to L.M.W., a joint University of Warwick-China Scholarship Council (CSC) PhD scholarship to M.L., European Research Council (ERC) awards to Y.C. (Grant agreement ID: 726116) and D.J.S. (Grant agreement ID: 883551), the NSF of China to C.-Y.L. (42276102), Y.-Z.Z. (32330001, W2441012) and N.W. (32200018), and the Taishan Scholars Program of Shandong Province, China to Y.-Z.Z. (tspd20240806) and C.-Y.L. (tsqn202306092).

Author affiliations: <sup>a</sup>MOE Key Laboratory of Evolution and Marine Biodiversity, State Key Laboratory of Marine Food Processing and Safety Control, College of Marine Life Sciences & Frontiers Science Center for Deep Ocean Multispheres and Earth System, Ocean University of China, Qingdao 266003, China; <sup>b</sup>State Key Laboratory of Microbial Technology, Marine Biotechnology Research Center, Shandong University, Qingdao 266237, China; <sup>c</sup>School of Life Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom; <sup>d</sup>School of Biosciences, University of Birmingham, Edgbaston B15 2TT, United Kingdom; <sup>e</sup>Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom; <sup>f</sup>Molecular Microbiology: Biochemistry to Disease, School of Biosciences, University of Sheffield, Sheffield S10 2TN, United Kingdom; and <sup>g</sup>Institute of Microbiology and Infection, University of Birmingham, Edgbaston B15 2TT, United Kingdom

Author contributions: Y.-Z.Z., D.J.S., and Y.C. designed research; N.W., L.M.W., M.L., C.-Y.L., A.R.J.M., Z.G., E.S., and I.D.E.A.L. performed research; N.W., L.M.W., M.L., C.-Y.L., A.R.J.M., Z.G., E.S., C.A.B., I.D.E.A.L., Y.-Z.Z., D.J.S., and Y.C. analyzed data; and N.W., L.M.W., M.L., C.-Y.L., C.A.B., I.D.E.A.L., Y.-Z.Z., D.J.S., and Y.C. wrote the paper.

1. Y. Koga, Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol. Mol. Biol. Rev.* **71**, 97–120 (2007).
2. L. Villanueva *et al.*, Bridging the membrane lipid divide: Bacteria of the FCB group superphylum have the potential to synthesize archaeal ether lipids. *ISME J.* **15**, 168–182 (2021).
3. D. M. Karl, Microbially mediated transformations of phosphorus in the sea: New views of an old cycle. *Ann. Rev. Mar. Sci.* **6**, 279–337 (2014).
4. J. I. Hedges *et al.*, Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature* **409**, 801–804 (2001).
5. R. Pedrosa-Pàmies, M. H. Conte, J. C. Weber, R. Johnson, Carbon cycling in the Sargasso Sea water column: Insights from lipid biomarkers in suspended particles. *Prog. Oceanogr.* **168**, 248–278 (2018).
6. L. M. Westermann *et al.*, Bacterial catabolism of membrane phospholipids links marine biogeochemical cycles. *Sci. Adv.* **9**, ead5122 (2023).
7. L. Behrendt *et al.*, Microbial dietary preference and interactions affect the export of lipids to the deep ocean. *Science* **385**, eaab2661 (2024).
8. S. Schouten, E. C. Hopmans, J. S. S. Damsté, The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. *Org. Geochem.* **54**, 19–61 (2013).
9. L. Xu, H. Guang, Z. Su, L. Liu, J. Tao, Diagenetic fate of glycerol ethers revealed by two novel isoprenoid hydroxyphytyl glycerol monoethers and non-isoprenoid alkyl glycerol ethers. *Org. Geochem.* **163**, 104344 (2022).
10. J. Logemann *et al.*, A laboratory experiment of intact polar lipid degradation in sandy sediments. *Biogeochem.* **8**, 2547–2560 (2011).
11. X. L. Liu *et al.*, From ether to acid: A plausible degradation pathway of glycerol dialkyl glycerol tetraethers. *Geochim. Cosmochim. Acta* **183**, 138–152 (2016).
12. B. Gašparović, A. Penezić, R. S. Lampitt, N. Sudasinghe, T. Schaub, Phospholipids as a component of the oceanic phosphorus cycle. *Mar. Chem.* **205**, 70–80 (2018).
13. J. J. Paraiso *et al.*, The distribution and abundance of archaeal tetraether lipids in U.S. Great Basin hot springs. *Front. Microbiol.* **4**, 247 (2013).
14. C. A. Francis, J. M. Beman, M. M. Kuypers, New processes and players in the nitrogen cycle: The microbial ecology of anaerobic and archaeal ammonia oxidation. *ISME J.* **1**, 19–27 (2007).
15. K. Knittel, A. Boetius, Anaerobic oxidation of methane: Progress with an unknown process. *Annu. Rev. Microbiol.* **63**, 311–334 (2009).
16. M. B. Karner, E. F. DeLong, D. M. Karl, Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**, 507–510 (2001).
17. T. Hoshino, F. Igagaki, Abundance and distribution of archaea in the seafloor sedimentary biosphere. *ISME J.* **13**, 227–231 (2018).
18. Y. M. Bar-On, R. Milo, The biomass composition of the oceans: A blueprint of our blue planet. *Cell* **179**, 1451–1454 (2019).
19. H.-C. Flemming, S. Wuerz, Bacteria and archaea on earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **17**, 247–260 (2019).
20. S. Muthusamy, F. Baltar, J. M. González, J. Pinhassi, Dynamics of metabolic activities and gene expression in the *Roseobacter* clade bacterium *Phaeobacter* sp. strain MED193 during growth with thiosulfate. *Appl. Environ. Microbiol.* **80**, 6933–6942 (2014).
21. S. Wuttge *et al.*, Determinants of substrate specificity and biochemical properties of the sn-glycerol-3-phosphate ATP binding cassette transporter (UgpB-AEC<sub>2</sub>) of *Escherichia coli*. *Mol. Microbiol.* **86**, 908–920 (2012).
22. Y. Koga, From promiscuity to the lipid divide: On the evolution of distinct members in Archaea and Bacteria. *J. Mol. Evol.* **78**, 234–242 (2014).
23. Y. Koga, N. Sone, S. Noguchi, H. Morii, Transfer of pro-R hydrogen from NADH to dihydroxyacetonephosphate by sn-glycerol-1-phosphate dehydrogenase from the archaeon *Methanothermobacter thermautotrophicus*. *Biosci. Biotechnol. Biochem.* **67**, 1605–1608 (2003).
24. J. Choe, D. Guerra, P. A. M. Michels, W. G. J. Hol, *Leishmania mexicana* glycerol-3-phosphate dehydrogenase showed conformational changes upon binding a bi-substrate adduct. *J. Mol. Biol.* **329**, 335–349 (2003).
25. V. Carbone *et al.*, Structure and evolution of the archaeal lipid synthesis enzyme sn-glycerol-1-phosphate dehydrogenase. *J. Biol. Chem.* **290**, 21690–21704 (2015).
26. R. P. A. Berntsson, S. H. J. Smits, L. Schmitt, D. J. Slotboom, B. Poolman, A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606–2617 (2010).
27. G. H. Scheepers, J. A. Lycklama A Nijeholt, B. Poolman, An updated structural classification of substrate-binding proteins. *FEBS Lett.* **590**, 4393–4410 (2016).
28. K. Fukami-Kobayashi, Y. Tateno, K. Nishikawa, Domain dislocation: A change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286**, 279–290 (1999).
29. T. Steiner, G. Koellner, Hydrogen bonds with pi-acceptors in proteins: Frequencies and role in stabilizing local 3D structures. *J. Mol. Biol.* **305**, 535–557 (2001).
30. C. Bisson *et al.*, The molecular basis of phosphite and hypophosphite recognition by ABC-transporters. *Nat. Commun.* **8**, 1746–1746 (2017).
31. I. Allice, E. R. Schreier, Crystal structure of *Escherichia coli* PhnD in complex with 2-aminoethyl phosphonate. *J. Mol. Biol.* **414**, 356–369 (2011).
32. M. Chandravanshi, R. Samanta, S. P. Kanaujia, Structural and thermodynamic insights into the novel dinucleotide-binding protein of the ABC transporter unveils its moonlighting function. *FEBS J.* **288**, 4614–4636 (2021).
33. J. S. Fenn *et al.*, 2019 Structural basis of glycerophosphodiester recognition by the *Mycobacterium tuberculosis* substrate-binding protein UgpB. *ACS Chem. Biol.* **14**, 1879–1887 (2019).
34. H. Shropshire, R. A. Jones, M. M. Aguilo-Ferretjans, D. J. Scanlan, Y. Chen, Proteomics insights into the *Burkholderia cenocepacia* phosphorus stress response. *Environ. Microbiol.* **23**, 5069–5086 (2021).
35. M. Sebastian *et al.*, Lipid remodelling is a widespread strategy in marine heterotrophic bacteria upon phosphorus deficiency. *ISME J.* **10**, 968–978 (2016).
36. J. J. Almagro Armenteros *et al.*, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
37. W. Minor, M. Cymborowski, Z. Otwinowski, M. Chruszcz, HKL-3000: The integration of data reduction and structure solution from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859–866 (2006).
38. M. D. Winn *et al.*, Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
39. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
40. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
41. P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
42. R. H. B. Smith, A. C. Dar, A. Schlessinger, PyVOL: A PyMOL plugin for visualization, comparison, and volume calculation of drug-binding sites. *bioRxiv* [Preprint] (2019). <https://doi.org/10.1101/816702> (Accessed 2 January 2025).
43. A. Milanese *et al.*, Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat. Commun.* **10**, 1014 (2019).
44. N. Wang *et al.*, 9LF9. Protein Data Bank. <https://doi.org/10.2210/pdb9LF9/pdb>. Deposited 8 January 2025.
45. N. Wang *et al.*, 9LFB. Protein Data Bank. <https://doi.org/10.2210/pdb9LFB/pdb>. Deposited 8 January 2025.
46. N. Wang *et al.*, 9LFF. Protein Data Bank. <https://doi.org/10.2210/pdb9LFF/pdb>. Deposited 8 January 2025.
47. N. Wang *et al.*, 9LEJ. Protein Data Bank. <https://doi.org/10.2210/pdb9LEJ/pdb>. Deposited 8 January 2025.