communications biology

Article

https://doi.org/10.1038/s42003-024-07258-3

A novel in-silico model explores LanM homologs among *Hyphomicrobium* spp

James J. Valdés [1,2] ✉, **Daniel A. Petrash** [3,4] ✉ & **Kurt O. Konhauser** [5]

Investigating microorganisms in metal-enriched environments holds the potential to revolutionize the sustainable recovery of critical metals such as lanthanides (Ln^{3+}). We observe *Hyphomicrobium* spp. as part of a Fe^{2+}/Mn^{2+}-oxidizing consortia native to the ferruginous bottom waters of a Ln^{3+}-enriched lake in Czechia. Notably, one species shows similarities to recently discovered bacteria expressing proteins with picomolar Ln^{3+} affinity. This finding was substantiated by developing an in-silico ionic competition model and recombinant expression of a homolog protein (*Hm*-LanM) from *Hyphomicrobium methylovorum*. Biochemical assays validate *Hm*-LanM preference for lighter Ln^{3+} ions (from lanthanum to gadolinium). This is comparable to established prototypes. Bioinformatics analyses further uncover additional *H. methylovorum* metabolic biomolecules in genomic proximity to *Hm*-LanM analogously dependent on Ln^{3+}, including an outer membrane receptor that binds Ln^{3+}-chelating siderophores. These combined observations underscore the remarkable strategy of *Hyphomicrobium* spp. for thriving in relatively Ln^{3+} enriched zones of metal-polluted environments.

Emerging clean technologies are driving the global demand for Ln^{3+}, also known as rare earth elements (REEs). Within the next decade, demand for Ln^{3+} is expected to increase ninefold, potentially surpassing known resources in 2035[1–3]. Currently, the separation, concentration, and purification of Ln^{3+} involve resource-intensive, energy-demanding metallurgical processes with significant environmental impact. China dominates the production and consumption of Ln^{3+}, supplying over 80% of the world's demand and controlling ~75% of the global market[3]. As the value of Ln^{3+} increases, so will their environmental impacts and possibility to fuel international conflicts[4,5]. Therefore, implementing innovative, eco-friendlier, and cost-effective alternatives for the extraction and recovery of Ln^{3+} is crucial to curb environmental degradation, ease geopolitical tensions, and secure a sustainable supply of these critical metals.

Biotechnological advances in the recovery of Ln^{3+} is foreshadowed by lanmodulin (LanM)[6–8], a periplasmic protein expressed by the Alphaproteobacteria *Methylobacterium extorquens* (*Mex*-LanM). *Mex*-LanM has a high affinity for Ln^{3+} at picomolar concentrations[9], demonstrating its strong lanthanophilic nature. LanM homologs have been identified among the nitrogen-fixing Alphaproteobacteria genus, *Bradyrhizobium*[9], and in the methylotroph *Beijerinckiaceae*, strain RH AL1[10]. Three physicochemical properties attributed to the apparent uniqueness of *Mex*-LanM[9] involve residues at specific EF-hand (EF1 – EF4) motif positions that stabilize binding (Supplementary Figs. S1 and S2 and Supplementary Table S1)[11] or contribute to the high selectivity of Ln^{3+} over calcium (Ca^{2+})[9] — Ca^{2+} being the native ion for most proteins with EF-hands[12]. Mutating these residues decreases the selectivity of Ln^{3+} over Ca^{2+} [9,11,13]. A fourth distinguishing property of *Mex*-LanM is the reduced number of intra-residues between EF-hands (~12 residues) compared to most proteins with EF-hands[9]. Gutenthaler et al. found that isolated *Mex*-LanM EF-hand peptides are capable of binding Ln^{3+} over Ca^{2+}, albeit at micromolar concentrations due to a lack of motif residues fully coordinating Ln^{3+} [14]. Therefore, the intra-residues between *Mex*-LanM EF-hands are considered critical for the affinity of Ln^{3+}. A recently discovered distant *Mex*-LanM homolog from *Hansschlegelia quercus* (*Hans*-LanM) forms dimers during the binding process of distinct Ln^{3+} [15].

The Ln^{3+} utilization and transport (*lut*) cluster refer to exclusive biomolecules involved in microbial metabolism of Ln^{3+} [16]. LanM is a genome marker for proximal genes encoded within the *lut*-cluster, such as the intracellular transport and/or trafficking of Ln^{3+} [16]. While microbes usually produce extracellular siderophores as iron scavengers, a select group of siderophores can potentially function as chelators Ln^{3+}, known as lanthanophores[9,16–18]. Recently discovered is the methylolanthanin (*mll*) cluster of *Methylobacterium aquaticum* responsible for biosynthesizing a lanthanophore for the complexation of Ln^{3+} [19]. Other studies have postulated on the *sbn* gene cluster in *Methylorubrum spp*. biosynthesizing an analog of the Fe-chelating, α-hydroxycarboxylate siderophore,

---

[1]Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czechia. [2]Centre Algatech, Institute of Microbiology, Czech Academy of Sciences, Třeboň, Czechia. [3]Department of Environmental Geochemistry and Biogeochemistry, Czech Geological Survey, Prague 5, Czechia. [4]Institute of Soil Biology and Biogeochemistry, Biology Centre, Czech Academy of Sciences, České Budějovice, Czechia. [5]Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB, Canada. ✉e-mail: valdes@paru.cas.cz; daniel.petras@geology.cz

staphyloferrin B[17]. At the front-line in *Methylorubrum* trafficking of Ln^{3+} is the outer membrane TonB-dependent transporter (TBDT). Bioinformatics and experimental analyses have identified the *sbn*-cluster as responsible for *Methylorubrum* TBDT-mediated Ln^{3+} trafficking[17].

The presence of REEs triggers a “Ln^{3+} switch” in some microbes that possess both Ln^{3+}- and Ca^{2+}-dependent proteins leading to an overexpression of the periplasmic dehydrogenase XoxF. The “Ln^{3+} switch” is mediated by TBDT, LanM, ABC transporters, and histidine kinases[9,16,20,21]. Microbial methanol oxidation via XoxF activation relies on Ln^{3+}-bound pyrroloquinoline quinone (PQQ)[16]. Additionally, in *Pseudomonas putida* KT2440, the “Ln^{3+} switch”[21] involves two distinct pathways regulated by two PQQ-dependent (alcohol/methanol) dehydrogenases whose expression inversely responds to the presence or absence of Ln^{3+} [21]. Both *sbn* siderophore and PQQ are independently known to stimulate microbial growth[17,22] and solubilize Ln^{3+} complexes[17,23]. However, not all microbes require lanthanophores if environmentally available Ln^{3+} are in dissolved form that can be passively absorbed or transported, such as through ion channels. For instance, the extremophilic *Methylacidiphilum fumariolicum* SolV, which inhabits volcanic mudpots with highly solubilized (ionic) concentrations of Ln^{3+}, most likely do not require the active acquisition of lanthanophores in TBDT-mediated Ln^{3+} trafficking[10,24].

Only 126 microbes, mainly gram-negative bacteria, are known to possess the five genes (encoding PqqA, PqqB, PqqC, PqqD, and PqqE) required for PQQ biosynthesis[25]. More than two decades ago, a *Hyphomicrobium* sp. was discovered that produces a high yield of extracellular PQQ[26]. As facultative methylotrophs, *Hyphomicrobium* spp. oxidize methanol to formaldehyde via a PQQ-dependent methanol dehydrogenase. They exhibit metabolic flexibility, thriving on a variety of carbon and energy sources, particularly methylated compounds. *Hyphomicrobium* spp. are capable of both aerobic and nitrate-driven organic respiration (denitrification). Some *Hyphomicrobium* spp. are putative Mn-oxidizers[27] since their distribution spans a range of subsurface, typically manganous habitats (Supplementary Table S2). Nonetheless, this metabolic capability remains to be proven.

The ferruginous bottom waters of the post-mining Lake Medard, NW Bohemia (Sokolov, Czechia) contain up to 31 µM of Fe^{2+} and Mn^{2+} [28]. This feature is due to an anoxic (O_2-depleted) groundwater influx carrying metals sourced from the weathered lake bedrock. This bedrock also includes a mined Ln^{3+}-enriched lignite seam[29]. Near the sediment-water interface, the solubility of Ln^{3+} is governed by complexation with Fe^{3+} and $Mn^{(3+,4+)}$ oxyhydroxide nanoparticles[29]. These nanoparticles cluster with organic matter at the oxic-anoxic boundary zone of the water column, forming aggregates above the lake’s redoxcline. Bacteria below this redoxcline interact with the reactive organomineral aggregates through reductive dissolution and electron exchange. In addition to aggregate reactivity, microbial lanthanophore trafficking mechanisms may enhance the solubility of Ln^{3+} in the water column.

While the role of *Hyphomicrobium* spp. as Mn^{2+} oxidizers remain unproven under the aforementioned conditions, their diverse metabolic capabilities, particularly in environments rich in methylated compounds, underline their ecological significance. These bacteria thrive in various habitats where methylotrophy is prevalent (Supplementary Table S2), highlighting their adaptability and potential important role in biogeochemical cycles involving carbon and transition metal respiration[30,31]. The environmental availability of Ln^{3+} enhances specific microbial metabolisms, in particular methylotrophy[32–34]. Incidentally, similar to *M. extorquens*, some *Hyphomicrobium* spp. express the gene cluster responsible for methanol oxidation[35]. Herein, our novel in-silico competition model for ionic binding reveals homologous *Hyphomicrobium* LanM proteins. Integrating the hydrochemistry and geomicrobiology of Lake Medard, our findings suggest that discrete *Hyphomicrobium* spp. possess an adaptable *lut*-cluster suited for environments enriched in Ln^{3+}.

Results

Hyphomicrobium spp. across the redoxcline

Building upon the biogeochemical features of Lake Medard bottom waters (Fig. 1A–D), the distribution of *Hyphomicrobium* spp. aligns with the unique redox gradients and dynamics of Ln^{3+} resulting from interactions between the redox-stratified water column, the planktonic microbiome inhabiting it, and the sediment-water interface. Within the microbial niche, a pronounced co-location of *Hyphomicrobium* spp. is observed in the microaerophilic zone above the redoxcline (Fig. 1E), where they constitute a substantial portion (0.9%) of the normalized microbiome. However, the relative abundance of *Hyphomicrobium* spp. sharply declines within this microaerophilic zone (≤0.2%), with a complete absence below 52 m of depth (Fig. 1E).

The distribution pattern of *Hyphomicrobium* spp. at Lake Medard corresponds with biologically mediated respiration processes that control the precipitation, mineralogical transformation, and dissolution of Fe(III) and Mn(III,IV) oxyhydroxides within the lake water column[28,29,36,37]. Complexation of Ln^{3+} (Supplementary Fig. S3) with insoluble iron oxyhydroxides formed in the chemocline may exert selective environmental pressure, particularly in microbial communities engaging in methylotrophy, driving microbial adaptation towards more efficient lanthanide uptake strategies[19,38]. A comprehensive analysis of the microbial amplicon dataset of Lake Medard[39] revealed 11 out of 12 unassigned/uncharacterized *Hyphomicrobium* spp. within this zone (Supplementary Figs. S4). Their phylogenetic placement forms distinct clades separate from well-characterized species. Notably, one *Hyphomicrobium* sp. (NCBI accession: MN664022) predominantly exists at 52 m, where Fe^{3+}- and Mn^{4+}-reduction becomes more prevalent (Fig. 1B). This distinct species is closely related to *H. album* and clusters with *H. methylovorum*, *H. denitrificans*, and *H. facile* (Fig. 1F, and Supplementary Fig. S4), which also inhabit sites enriched with Ln^{3+} (Supplementary Table S2). Below 52 m, marking a notable shift in the lake’s geochemical environment (i.e., the hypolimnion-monimolimnion interface), peak dissolved concentrations of Ln^{3+} ($\Sigma[REE] = 711$ pM) are recorded within the anoxic waters (Fig. 1C, and Supplementary Table S3).

A collection of putative *Hyphomicrobium* LanM proteins

Recent bioinformatics by Cotruvo and co-workers revealed 696 *Mex*-LanM homologs—a sub-group contained remote homologs, including *Hans*-LanM[15]. Herein, a protein PHI-BLAST[40] against the entire non-redundant NCBI database, solely based on the EF-hand Ln^{3+} binding motif (Supplementary Fig. S1), resulted in 298 protein sequences with high similarity to *Mex*-LanM (49–70%). After selecting proteins that express a signal peptide and removing duplicates (>90% identical), 52 sequences predominantly of Alphaproteobacteria origin were retained (Supplementary Table S4). The most represented Alphaproteobacteria families (genera) include *Nitrobacteraceae* (*Bradyrhizobium*, *Tardiphaga*, and *Rhodopseudomonas*) and *Hyphomicrobiaceae* (*Hyphomicrobium, Methyloligella*). *Streptomyces purpurogeneiscleroticus*, a member of the *Actinomycetia* class (family *Streptomycetaceae*), was also identified. Bacterial genera such as *Hyphomicrobium*, *Beijerinckia*, *Methylocystis* and *Streptomyces* (Supplementary Table S4) particularly thrive in the bottom-water column of Lake Medard (Fig. 1F and Supplementary Fig. S4)[28,41].

All 52 proteobacterial LanM homologs we identified feature four EF-hands (Fig. 2A), with short intra-residues and one to three D_1-N_1 substitution(s) at the first motif position (Fig. 2A). These are among the attributes that contribute to the unique Ln^{3+} binding properties of LanM[9]. Although motif positions D_3 and E_{12} are highly conserved in all 52 LanM homologs, approximately half of the homologs vary in motif residue composition. Specifically, the homologs contain residue substitutions at positions D_1, P_2, D_5, T_7 and D_9 (Fig. 2A). These substitutions, specifically at motif positions D_1, P_2, and D_9, are attributed to a higher preference for light rather than heavy REEs (LREEs *vs.* HREEs)—as in the remote LanM homolog from *H. quercus* (*Hans*-LanM). The prototypal *Mex*-LanM (Supplementary Fig. S1) preference for REEs is not specific[9,11,15].

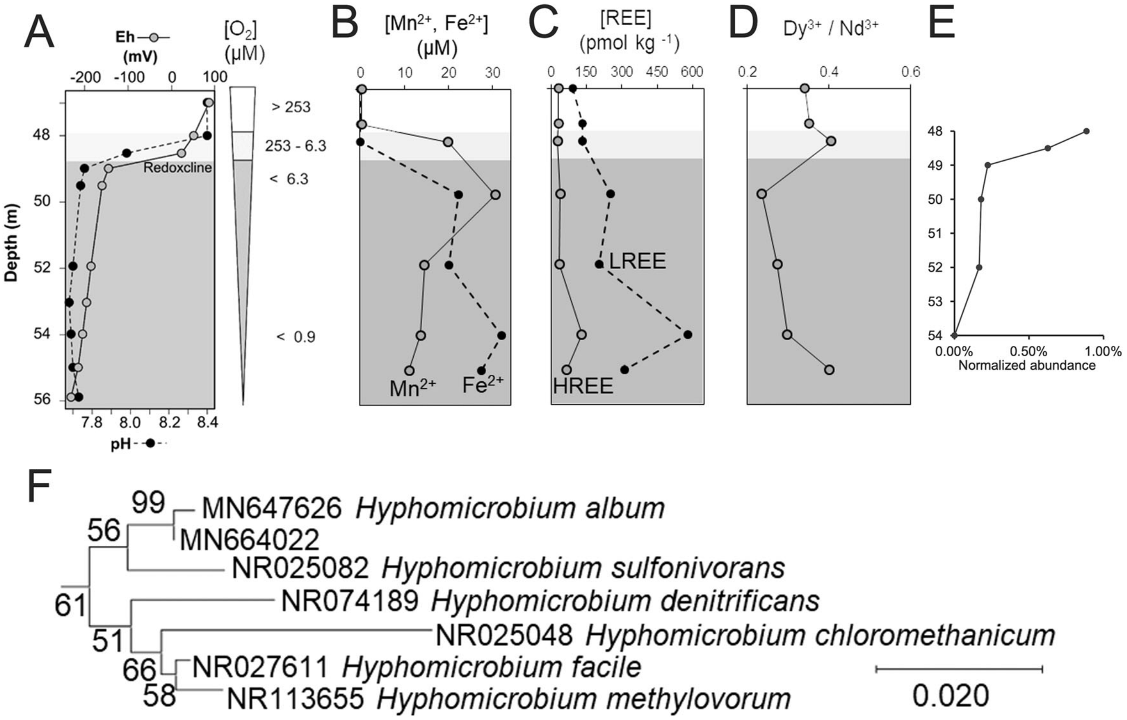


Fig. 1 | The hydrochemistry of the bottom-water column of Lake Medard. A The strongly redox stratified water column is characterized by circumneutral pH. The redoxcline (i.e., inflection point in the Eh curve) is shown. **B** In the O$_2$-depleted water column, dissolved Mn^{2+} and Fe^{2+} concentrations vary with depth[28], while light (LREE: lanthanum (La) through Europium, Eu) and heavy REE (HREE: gadolinium (Gd) through lutetium (Lu)) concentrations follows the trend imposed by particulate Fe^{3+} and Mn^{3+} dissolution. **C** The sum of dissolved LREEs and HREEs concentrations sourced from the lake bedrock, coal post-mining spoils, and underlying Miocene rift lake sediments. **D** The molar ratio of Dy^{3+} and Nd^{3+}, as a proxies for LREE:HREE partitioning, respectively, in the O$_2$-depleted Lake Medard water column. **E** The normalized abundance of *Hyphomicrobium* spp. (*y*-axis) indigenous to the Lake Medard (LM) bottom-water column (*x*-axis). Data was collated from the LM microbial amplicon dataset[39]. **F** A clade from the *Hyphomicrobium*-based 16S rRNA neighbor-joining phylogenetic tree, with outgroup (Supplementary Fig. S4), showing the top 4 classified *Hyphomicrobium* spp. possessing LanM homologs (Fig. 2, and Supplementary Figs. S5) and the Lake Medard-dwelling *Hyphomicrobium* spp. (NCBI accession number only) related to *H. album*. The scale bar (bottom right corner) and bootstrap values ≥50% are shown.

A protein-based phylogram based on the EF-hand motifs represents homologs among *Beijerinckia*, *Hyphomicrobium Methyloligella*, *Methylocystis* and *Streptomyces* clustering with LanM, with *Hyphomicrobium* (11 species) being the predominant genus (Fig. 2B and Supplementary Fig. S5). Accordingly, the primary and tertiary structures of four classified *Hyphomicrobium* spp. were compared to *Mex*-LanM (Supplementary Fig. S6). *Mex*-LanM EF-hand primary sequence conservation ranges, in descending order, as EF2 > EF3 > EF4 > EF1. Among these homologs, the *H. facile* sequence reveal four EF-hand motifs without the D$_1$-N$_1$ substitution. All four *Hyphomicrobium*-predicted tertiary structures are strikingly similar to *Mex*-LanM (Fig. 2C).

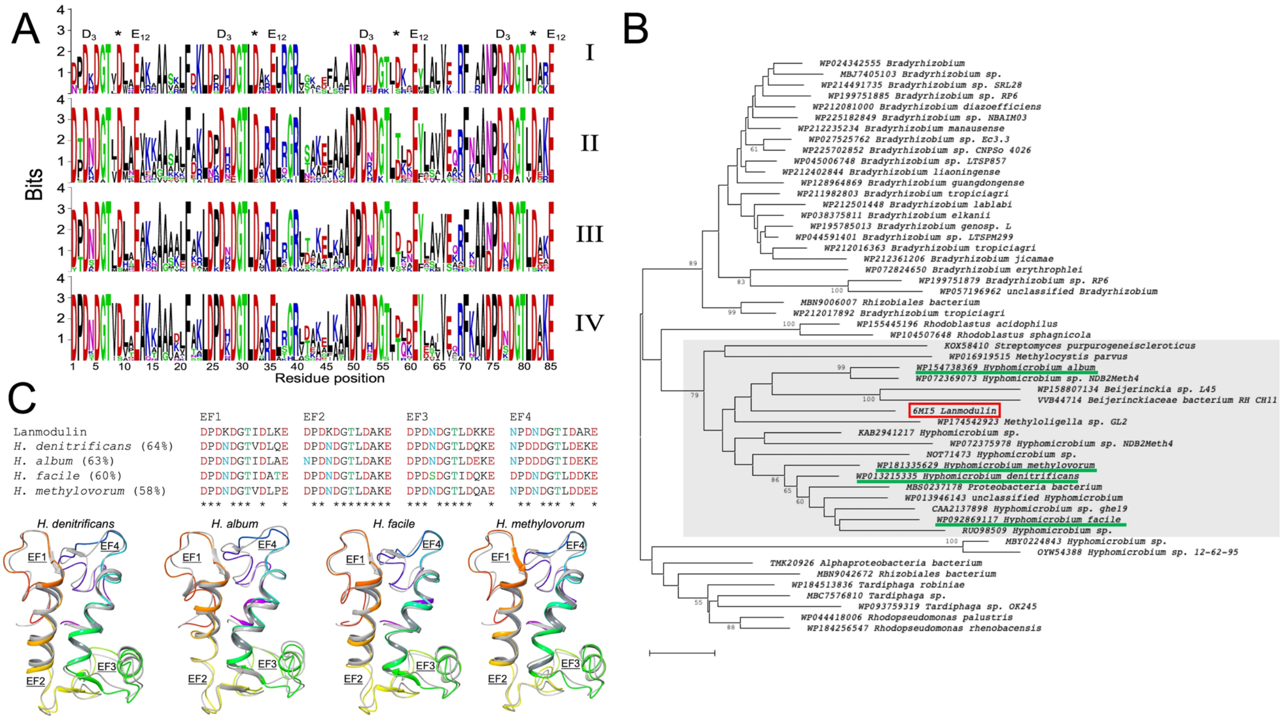
The *Hyphomicrobium lut*-cluster for metabolizing REEs

Phylogenetically relevant to the *Hyphomicrobium* sp. at 52 m depth in Lake Medard (Fig. 1F) and other model sites (Supplementary Table S2), the NCBI genomes of *H. album* and *H. methylovorum* were analyzed. The methanol/ethanol PQQ-dependent dehydrogenase genes are located 2.5 kbp downstream and 11.3 kbp upstream to their respective LanM encoding genes (Fig. 3). The predicted structures of these *Hyphomicrobium* PQQ-dependent (alcohol/methanol) dehydrogenase are comparable to the resolved XoxF, a Ln^{3+}-dependent PQQ-dependent (alcohol/methanol) dehydrogenase found in the periplasmic space of *Methylomicrobium buryatense*[20], with an α-carbon backbone deviation of ~1.2 Å (Supplementary Fig. S6). *H. album* has a single Asp-Ala substitution within the dehydrogenase active site (Supplementary Fig. S6), which may reduce PQQ binding efficiency[42]. However, these *Hyphomicrobium* PQQ-dependent (alcohol/methanol) dehydrogenases are part of the gene cluster encoding dehydrogenases *moxJ* and *moxG* that are dependent on Ca^{2+}, unlike the XoxF-type, REE-dependent counterparts[43].

Upon detailed inspection of both genomes, *H. methylovorum* encodes an *moxF-xoxF* gene 5.8 kbp downstream of its LanM gene, which is absent in *H. album* (Fig. 3). The closest homolog found during protein structure prediction of the *H. methylovorum moxF-xoxF* gene product (*Hm*XoxF) was the XoxF1-type methanol dehydrogenase, bound to PQQ complexed with the REE neodymium, from *Methylacidimicrobium* sp. AP8[44]. *Hm*XoxF exhibits 62% sequence similarity to XoxF1, and the predicted structure of *Hm*XoxF shows an α-carbon backbone deviation of 0.4 Å compared to XoxF1 (Fig. 3, inset). The active site of XoxF1, complexed with PQQ-neodymium, is highly conserved in *Hm*XoxF (Fig. 3, inset).

BLAST analysis confirms that *M. extorquens* possess the *sbnA* – *sbnH* encoding genes responsible for synthesizing siderophore analogs of staphyloferrin B (SE8)[17]. On the one hand, *H. album* and *H. methylovorum sbnH* were identified, while *sbnD* (*H. methylovorum*) and *sbnA* (*H. album*) are also present. On the other hand, these *Hyphomicrobium* spp. are equipped with the obligatory encoding genes for extracellular PQQ production. The percent sequence conservation for these findings are indicated (Supplementary Tables S5 and S6). Notably, the *sbnH* gene product homologs in *Hyphomicrobium* are diaminopimelate decarboxylases, which are required for microbial hydroxamate siderophore biosynthesis, such as ferrioxamines[17,45], and may be transported by TBDTs.

The genomes of *M. extorquens*, *H. methylovorum*, and *H. album* encode 22–24 TBDT genes. Of these TBDTs, 15 from *M. extorquens*, 9 from *H. album*, and 8 from *H. methylovorum* are >600 amino acids long and contain a signal peptide. In contrast to the gene encoding the *M. extorquens* TBDT (*Me*TBDT), which is adjacent to the *Mex*-LanM gene[16], the *H. methylovorum* and *H. album* TBDT genes are slightly farther from their respective LanM protein-encoding gene (Fig. 3). The *H. methylovorum* TBDT is 19.7 kbp downstream of its LanM gene, with the *H. album* being



	EF1	EF2	EF3	EF4
Lanmodulin	DPDKDGTIDLKE	DPDKDGTLDAKE	DPDNDGTLDKKE	NPDNDGTIDARE
H. denitrificans (64%)	DPDNDGTVDLQE	DPDNDGTLDAKE	DPDNDGTLDQKE	NPDDDGTLDEKE
H. album (63%)	DPDNDGTIDLAE	NPDNDGTLDAKE	DPDNDGTLDEKE	NPDDDGTIDEKE
H. facile (60%)	DPDNDGTIDATE	DPDNDGTLDAKE	DPDSDGTIDQKE	DPDNDGTLDDKE
H. methylovorum (58%)	DPDNDGTVDLPE	DPDNDGTLDAKE	DPDNDGTLDQAE	NPDNDGTLDDEE
	`*** *** *  *`	` ** ********`	`*** *****  *`	` ** *** *  *`

Fig. 2 | LanM homologs. A Web-logo[85] generated graph-based on the EF-hand loops of the *Mex*-LanM homologs ($n = 52$). The conserved D_3 and E_{12} of each EF-hand loop are labeled; * = position 9. The Roman numerals distinguish whether the homologs express one, two, three or four EF-hand loops with a conserved motif position D_1. **B** Neighbor-joining phylogram based on the EF-hand loop sequences of the 52 identified proteobacteria homologues that describe similarities with LanM (red outline). The *Mex*-LanM sub-clade is the shaded region. *Hyphomicrobium* spp. That were additionally computationally analyzed are underlined in green. The branch length scale bar is 0.05 and bootstrap values above 50% are shown. **C**, above The primary sequences of each *Hyphomicrobium* EF-hand with percent similarity (parenthesis) compared to *Mex*-LanM indicated. The asterisks denote conserved residues. **C**, below The *Hyphomicrobium* tertiary structures color-traced from the amine-terminus (red) to the carboxy-terminus (purple) structurally aligned with *Mex*-LanM (grey). The EF-hands are labeled.

flanked by two TBDTs, 12.6 kbp upstream and 19.7 kbp downstream (Fig. 3 and Supplementary Fig. S7).

In-silico *Me*TBDT and *H. methylovorum* TBDT (*Hm*TBDT) siderophore binding was investigated using α-hydroxycarboxylates, catecholates, and hydroxamates (Supplementary Results and Supplementary Figs. S8–S11). Siderophore binding with *H. album* TBDTs were undetermined due to its in-silico structural instability (Supplementary Results and Supplementary Figs. S7 and S8). The *Me*TBDT siderophore binding and MD simulations corroborate the lanthanophore analog experiments regarding *Methylobacterium sbn* (Supplementary Results and Supplementary Figs. S8 and S9)[17]. Ferrioxamine G (FOG) was the top hydroxamate siderophore for *Hm*TBDT, with favorable binding scores and active site conformations (Supplementary Results and Supplementary Fig. S10). A simulation of 500-ns MD on *Hm*TBDT-FOG presents the active site interacting residues that are conserved upon TBDT-siderophore binding (Supplementary Results and Supplementary Figs. S10 and S11).

Mex-LanM *vs.* select *Hyphomicrobium* LanM structures during ionic competition

Molecular dynamics (MD) simulations can model nano-events pertaining to LanM binding using recently established trivalent metal (i.e., Ln^{3+}) force fields, allowing for the estimation of atomic distances and free energies consistent with experimental results[46–48]. The nuclear magnetic resonance (NMR) structures of *Mex*-LanM (Fig. 4A)[11] and calbindin from *Rattus norvegicus*[49] serve as MD benchmarks for the in-silico ionic competition model (Supplementary Results and Supplementary Figs. S12–S14), preceding the assessment of *Hyphomicrobium* LanM homologs (Fig. 2C). Substitution at the first motif position (D_1-N_1) of EF4 in *Mex*-LanM causes a lower Ln^{3+} binding affinity, shifting from picomolar to micromolar concentrations[9,11]. This lower affinity for Ln^{3+} results in the absence of a fourth ion in the *Mex*-LanM resolved structure (Fig. 4A and Supplementary Fig. S1). Despite this, the EF4-Ln^{3+} affinity remains higher than the LanM-Ca^{2+} binding affinity, which is in the millimolar range[9,11]. Four *Mex*-LanM replicates illustrate ionic preference for Ln^{3+} over Ca^{2+}, maintaining EF4 distances 5–9 Å throughout the simulated competition (Supplementary Fig. S12A, S12B). The ion positions for each replicate were swapped and these subsequent simulations reveal that 11 replicates (out of 12) show high preference for Ln^{3+} over Ca^{2+}, confirming the validity of our in-silico competition model (Supplementary Fig. S12C–S12D). The MD simulations show Ln^{3+} occupying the EF4 site, while Ca^{2+} remains exposed to the solvent, away from the *Mex*-LanM structure (Supplementary Fig. S13).

Calbindin EF1 shows the highest affinity for Ln^{3+} among its EF-hands, while EF2 and EF6 lack ion binding capability[13,50]. Therefore, calbindin EF1 and EF5 serve as additional qualitative positive controls, and their adjacent loops, EF2 and EF6, as negative controls (Supplementary Fig. S14) for benchmarking our in-silico ionic competition model. As in the *Mex*-LanM EF4 ionic competition (Supplementary Fig. S12), only 1/3 of replicates show Ln^{3+} approximating the lower affinity calbindin EF5 (Supplementary Fig. S14). By contrast, the higher affinity EF1 of calbindin binds Ln^{3+}, instead of Ca^{2+}, in 1/2 of the replicates. Additionally, there was only a 15% margin of error for binding the negative control EF-hands, EF2 and EF6 (Supplementary Fig. S14), that were apparent artifacts during the MD simulations (Supplementary Results).

Yttrium (Y^{3+}), a chemically similar REE, was used to exclusively test *Mex*-LanM EF4 binding in the in-silico model. Three 20-ns MD replicates were simulated for *Mex*-LanM EF4-Y^{3+} binding using the lowest energy NMR conformation (Fig. 4A, B)[11]. In all three replicates, the Y^{3+} ion initially approaches 5~70 Å from the *Mex*-LanM EF4 until forming monodentate contacts with residue E119 (E_{12}) in within 10 ns (Fig. 4B). Additional Y^{3+} contacts form with EF4 residue D110 (D_3), while contacts with E119 begin alternating in denticity, mainly forming bidentate bonds at ~2.8 Å (Supplementary Table S7). The replicates were then simulated to complete 1-μs

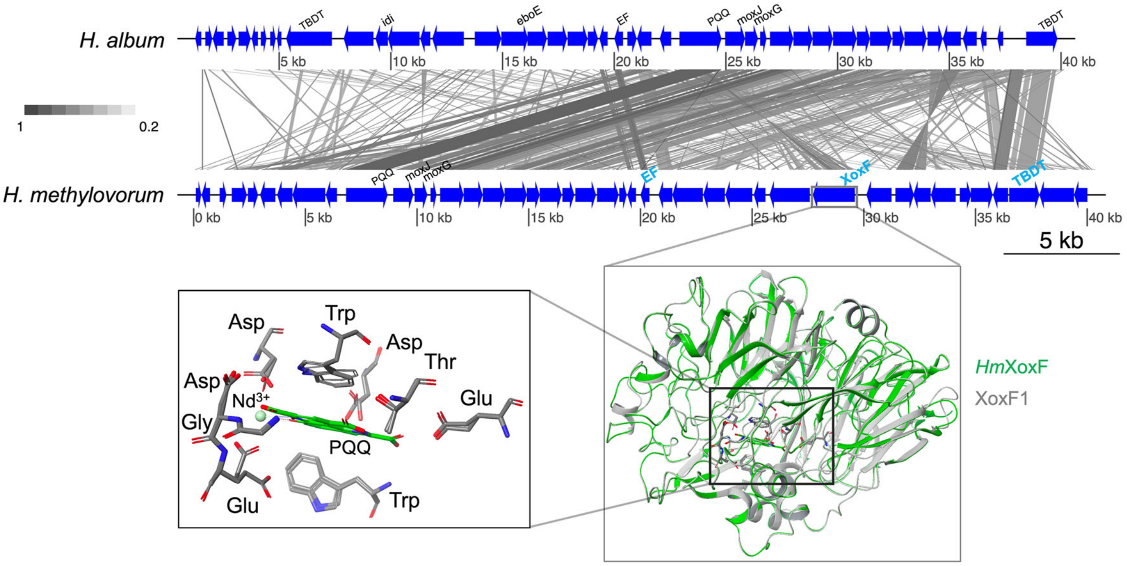


Fig. 3 | The *Hyphomicrobium lut*-cluster. The regional genomic schematic generated using the R-based genoPlotR package[84] represents the *H. methylovorum* (NZ_QHJE00000000.1) bidirectional genes within 20 kbp of the LanM protein encoding gene (EF). The matrix depicts the parallel, genomic relationship with *H. album* (NZ_WMBQ00000000.1). *H. methylovorum lut*-cluster are in cyan lettering. The scale bar (left-middle) indicates the level of gene homology. (Inset) The structural alignment is based on the predicted *Hm*XoxF (green) and XoxF1 PQQ-dependent (alcohol/methanol) dehydrogenase structures (grey; PDB: 7O6Z). The active site (left inset) depicts the conserved XoxF1 residues coordinating neodymium-PQQ.

MD (Supplementary Results) for EF4 to fully coordinate Y^{3+} (Fig. 4C). While fully coordinated, each EF-hand maintains the P_2-D_3 hydrogen bond that stabilizes the binding of Ln^{3+} (Supplementary Fig. S1) with average distances of ~2.5 Å (Supplementary Table S8). Interactions between the *Mex*-LanM N-terminus and the loop EF4, which influence motif residue coordination and water occupancy (Supplementary Figs. S13 and S15), accounts for its lower affinity[9,11,14].

During ionic competition, *Hyphomicrobium* EF1–EF3 maintained bound Y^{3+} ions, with an additional Y^{3+} and Ca^{2+} positioned equidistant from the EF4 (~15 Å), similar to the *Mex*-LanM MD simulations (Supplementary Fig. S12). Among the Mex-LanM homologs, those from *H. methylovorum* and *H. album* performed best, with 4 out of 5 replicates showing a preference for Y^{3+} binding over Ca^{2+} at EF4 (Fig. 4D). The Y^{3+} distances for these *Hyphomicrobium* LanM proteins (Fig. 4D), specifically *H. methylovorum*, approximate <5 Å toward EF4. Comparatively, the distances for *Mex*-LanM during ionic competition are higher (≥5 Å; Supplementary Fig. S12). The two remaining *Hyphomicrobium* LanM homologs maintain further EF4-Y^{3+} distances, but still mostly outperforming Ca^{2+} (Fig. 4D).

The lanthophilic protein of *H. methylovorum* shows high preference for LREEs

The lanthophilic protein of *H. methylovorum*, designated *Hm*-LanM, exhibits higher homology to *Mex*-LanM (58%) compared to *Hans*-LanM (33%). This suggests that *Hm*-LanM could be highly crucial in Ln^{3+} metabolism[16] at relevant environmental sites (Figs. 1 and 2, and Supplementary Table S2). We hypothesize that *Hm*-LanM exhibits a high preference for REEs over Ca^{2+}. To test this, the *Hm*-LanM gene was expressed in *Escherichia coli* Rosetta (DE3) cells and protein binding was examined with La^{3+}, europium (Eu^{3+}), and ytterbium (Yb^{3+}) as respective representatives of more reactive LREE (i.e., have no paired electrons in their 4*f* orbitals, and with atomic numbers 57 to 64), middle REE (MREE, Sm to Ho which exhibit intermediate electron configurations), and HREE (Gd to Lu), which have paired 4*f* electrons, and are less reactive than LREEs. Circular dichroism (CD) spectral analyses indicate that the recombinant *Hm*-LanM undergoes distinct conformational shifts at 222 nm under subsaturated concentrations of REEs. Interactions with MREE and LREE form α-helical structures, while interaction with HREE results in a less defined helix (Fig. 5A)[9].

The molecular weight (MW) of the unbound *Hm*-LanM apo-structure is 11.5 kDa (Supplementary Fig. S16), with an estimated isoelectric point of 4.7—both properties similar to *Mex*-LanM[9]. The REE-bound *Mex*-LanM is strictly a monomer at a MW ~ 15 kDa[9,15]. We used an 0.05% sodium lauroyl sarcosinate (sarkosyl) SDS-PAGE as a time-efficient and cost-effective method to detect dimer formation[51]. Gel electrophoresis determines that *Hm*-LanM approximates a higher MW upon REE exposure. Incubating *Hm*-LanM with LREE results in a formation at 27.5 kDa, while an additional lower MW band occurs in the presence of MREE and HREE (Supplementary Fig. S17). Treatment with Ca^{2+} strictly produces this lower MW band determined at 22.1 kDa. Considering that chelators compete for metal binding, high concentrations of EDTA with *Hm*-LanM also produces a prominent lower MW form.

Hm-LanM (8.7 μM), when saturated with LREE (1 mM), shows exceptional resilience and stability by maintaining the 27.5 kDa structure (Supplementary Fig. S17) despite extreme urea concentrations (920:1 urea to protein ratio), which often causes protein denaturation (Fig. 5B). Considering that chelators compete for metal binding, high concentrations of EDTA (10:1 EDTA:LREE) partially degraded the *Hm*-LanM complexed with LREE as evidenced by both MW forms being present (Fig. 5B). A decrease in MW was triggered only by combining EDTA and urea (Fig. 5B), indicating that LREE incorporation enhances the structural integrity of *Hm*-LanM. The band formation at 22.1 kDa, caused by EDTA, thus corresponds to the *Hm*-LanM apoprotein, while the 27.5 kDa MW is an REE-bound, monomer structure (Fig. 5B and Supplementary Fig. S17).

At lower REE concentrations, *Hm*-LanM (8.7 μM) forms a 27.5 kDa structure at ~5 equivalents (equiv.) of LREE (Fig. 5C and Supplementary Fig. S17). Indeed, the larger *Hm*-LanM form appears to occur with higher MREE concentrations (23 equiv.), but not with HREE. Consistent exposure to Ca^{2+} results in the formation of the apoprotein version of *Hm*-LanM (Fig. 5C and Supplementary Fig. S17). These prevalent protein complexation patterns reinforce the natural binding preference of *Hm*-LanM, with LREE > MREE > HREE over Ca^{2+}.

Discussion

Our study demonstrates that trivalent MD force fields[46–48] are effective for simulating ionic preference between Ca^{2+} or Ln^{3+} for LanM proteins (Fig. 4, and Supplementary Figs. S12–S15). Conventional MD protocols

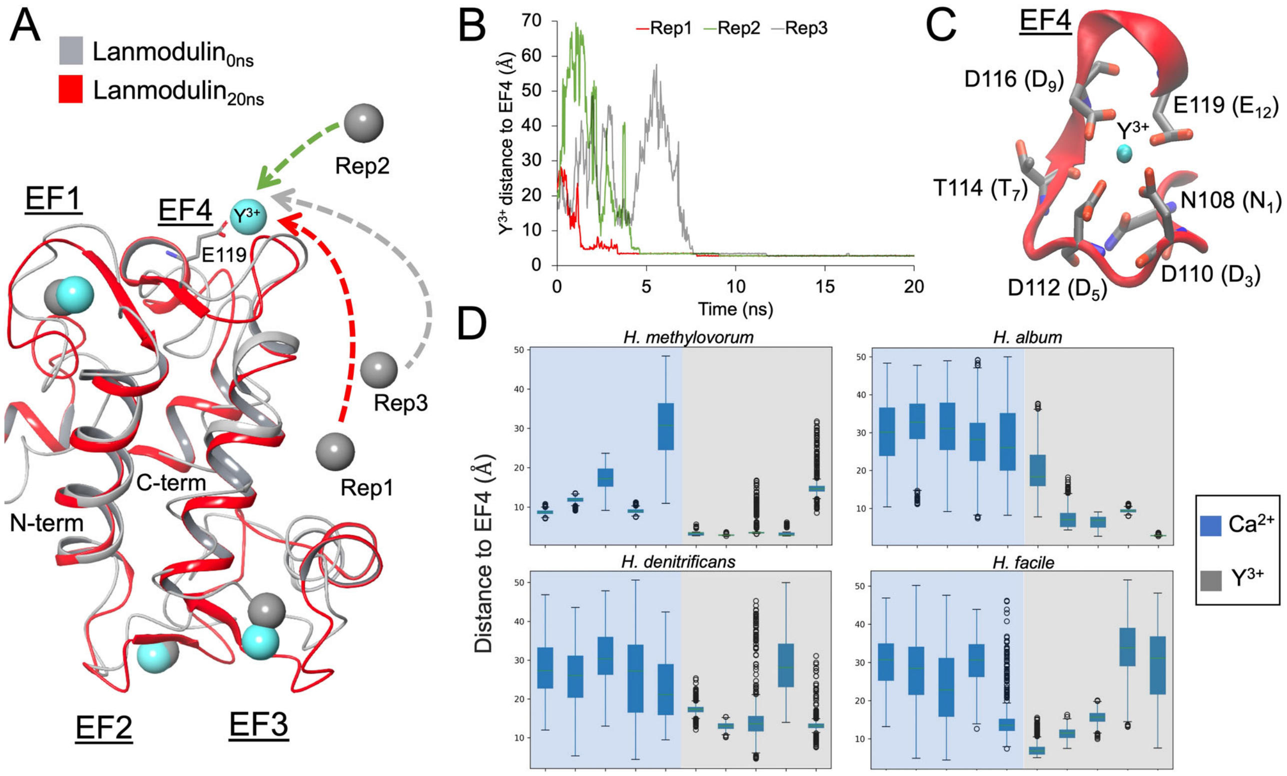


Fig. 4 | *Mex*-LanM EF4 binding and *Hyphomicrobium* ionic competition. **A** The superimposed *Mex*-LanM starting pose (grey) and end conformation (red) after 20-ns MD. The amine-terminus (N-term) and carboxy-terminus (C-term), EF-hands and E119 are labeled. The grey-cyan Y^{3+} spheres are MD starting-end positions, respectively. Dashed arrows mimic each Y^{3+} approach to EF4 (replicate spheres are numbered). **B** The distance for each Y^{3+} replicate (Rep) to EF4 (*y*-axis) during 20-ns MD (*x*-axis). **C** Birdseye snapshot of *Mex*-LanM EF4 motif residues (labeled) coordinating Y^{3+} after 1-μs MD. **D** Boxplots representing the ionic competition replicates (*x*-axis) for each *Hyphomicrobium* spp. during 20-ns MD.

have yet to explore ionic competition in this context. Furthermore, we hypothesize that the increased distances of Y^{3+} from EF4-N108 and EF4-T114 explains the lower affinity for Ln^{3+} compared to other LanM EF-hands (Supplementary Results and Supplementary Table S7)[9,11]. This lower affinity for EF4 binding is potentially dependent on *Mex*-LanM N-terminal contacts for Ln^{3+}, with simulations maintaining a hydration of 2–6 water molecules (Supplementary Results, and Supplementary Figs. S13 and S15), similar to observations in low-affinity *Mex*-LanM EF-hand peptides[14]. Our benchmarked ionic competition model and REE binding experiments also indicates that *Hm*-LanM produces a MW similar to the *Hans*-LanM dimer under LREE treatment with a moderate MREE and lower HREE preference (Figs. 4 and 5 and Supplementary Figs. S16). This data provides a compelling foundation to further investigate the structural dynamics and REE-binding properties of *Hm*-LanM as a robust tool for biotechnology, capable of withstanding complex biochemical conditions (Fig. 5). Additionally, as a member of consortia in REE-enriched environments (Supplementary Table S2), *H. methylovorum* has a sufficient *lut*-cluster for trafficking Ln^{3+}, evidenced by possessing key homologs in proximity to the *Hm*-LanM encoding gene, such as TBDT and XoxF[16] (Fig. 5, and Supplementary Figs. S5 and S16).

In-silico analyses of bacterial receptor binding combined with active site residues using MD have significantly advanced our understanding of microbial resistance to environmental pressures[52]. The homologous FoxA binding properties of *Me*TBDT, which coordinate initial electrostatic interactions and pi-stacking, are essential for TBDT-siderophore scavenging[53]. Although most *Hyphomicrobium* siderophore-binding results remain unresolved (Supplementary Figs. S8–S10), they do indicate that *Hm*TBDT is capable of coordinating ferrioxamine, FOG (Supplementary Figs. S10 and S11). Likewise, *Hyphomicrobium* spp. are equipped with *sbnH* (Supplementary Tables S5 and S6) that is required for microbial ferrioxamine biosynthesis[17,45]. Chemical analyses indicate that ferrioxamines can also function as lanthanophores, forming stable complexes to enhance trivalent ionic mobility by increasing Ln^{3+} solubility[54].

The solubility of Ln^{3+} is low at the dysoxic portion of the Lake Medard bottom-water column because of various factors, including pH, the presence of reactive ionic species such as (bi)carbonate and sulfate, Fe, and humic and fulvic acids[28]. These species can complex with Ln^{3+} to influence their solubility and mobility in aqueous solution (Fig. 1B–F, and Supplementary Figs. S3 and S4). The concentrations of Ln^{3+} in Lake Medard are comparable to those found in deep seawater[55], and the sediments exhibit Ln^{3+} within the range observed in deep-sea sediments[29]. The organomineral aggregates at Lake Medard, comprised of oxyhydroxide nanoparticles and microbial cells[28], influence the speciation of Ln^{3+} and other metals in solution[29]. These aggregates have a small electronegative surface area which confers them with a high sorption capacity toward Ln^{3+} at circumneutral pH[56]. The aggregates encompass a niche suitable for microbes capable of efficiently respiring particulate Mn and Fe stocks under anoxic conditions[28,57], with *Hyphomicrobium* spp. being particularly well-suited for this niche (Fig. 1, and Supplementary Figs. S3 and S4 and Table S2).

Microbial siderophore secretion for Fe-scavenging might offer an evolutionary advantage[58]. For example, microbes secreting high-affinity siderophores under Fe-limited settings will outcompete those with low-affinity siderophores. This competitive advantage is heightened for microbes that can scavenge and biosynthesize a diverse range of siderophores, dominating those that secrete fewer analogs. Cumulative research implicates this evolutionary arms race as a driving force for specific heavy-metal resistant microbes capable of thriving in highly polluted environments[58]. Our data unequivocally establishes that the *Hyphomicrobium lut*-cluster enables these microbes to effectively counteract a scarcity of Ln^{3+} at ferruginous redox interfaces (Figs. 1–4, Supplementary Figs. S3–S16, and Supplementary Tables S2–S8). The distribution of *Hyphomicrobium*

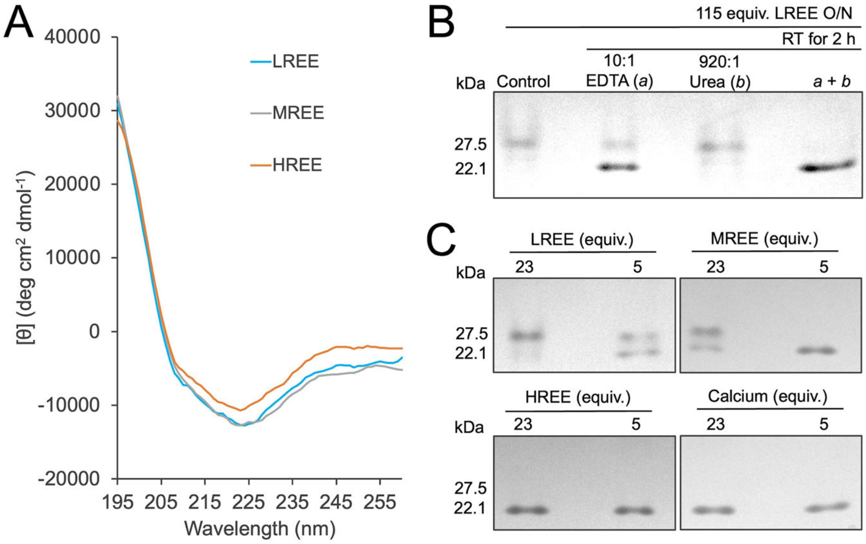


Fig. 5 | *Hm*-LanM CD spectra and gel mobility analysis. A The *Hm*-LanM (20 µM) CD spectra collected and treated with 1 µM REEs (or 0.05 equiv.). Both buffer sets were treated. **B** The purified *Hm*-LanM (8.7 µM) was incubated with 1 mM LREE (115 equiv.) overnight (i.e., O/N) and then treated for 2 h at room temperature (RT) with 10 mM EDTA (*a*), 8 M Urea (*b*), or a combination of both (*a* + *b*). The EDTA (or urea) to LREE (or protein) ratio is indicated. **C** The *Hm*-LanM (8.7 µM) was incubated with 200 or 40 µM REEs (equiv. specified respectively). The original gels are presented in Supplementary Fig. 18. Note: *Hm*-LanM precipitation during gel electrophoresis weaken the gel band resolution.

spp. at the study site, as in other Mn-enriched chemically stratified settings (Supplementary Table S2), likely reflects their adaption to Mn-oxyhydroxides, which are primary carriers of Ln^{3+} [59]. Consequently, aqueous environments where these phases are present create a chemical gradient associated with reduction processes that solubilize Ln^{3+}. This possibly enhances the ability of *Hyphomicrobium* to conduct methylotrophy using their robust *lut*-cluster, leading to their common interpretation as bacteria linked to the deposition of iron and manganese oxides[60].

Extending this study to other lanthanide-enriched environments (Supplementary Table S2), including additional post-mining lakes, can provide a broader understanding of convergent evolutionary trends in the respiration of metals and methylated-compounds at aqueous redox interfaces These future analyses could further support our findings and strengthen the hypothesis that *Hyphomicrobium* spp. utilize Ln^{3+} without necessarily involving an active Mn^{2+} or Fe^{2+} oxidation functioning.

Methods

Lake Medard hydrochemistry

For the Lake Medard hydrochemistry, a water quality monitoring and profiling probe (YSI 6600 V2-2) was used for measuring O_2 concentrations, pH and Eh in the stratified portion of the Lake Medard bottom-water column. Bottom-water sampling was implemented by using a Ruttner sampler. Dissolved cations concentrations were determined in filtered (0.45 µM cutoff), acidified water sample aliquots using a Thermo Element 2, high-resolution inductively coupled plasma mass spectrometer (HR-ICP-MS). For equilibrium modeling anion concentrations were also determined in filtered un-acidified aliquots analyzed by high-pressure liquid chromatography (HP-LC). The concentration data were calibrated against multi-element standards at variable concentrations that were repeatedly measured throughout the analytical session[28].

Sequence alignment and phylogenetic analysis

Alignments were created and edited by truncating flanking sequences using the MEGA software[61]. The evolutionary history was inferred using the Neighbor-Joining (NJ)[62] and Maximum Likelihood method (ML) methods. The percentage of replicate trees associated with sequences clustered together in the bootstrap test (1000 replicates)[63]. The distances were computed using the Kimura 2-parameter method[64] with the number of substitutions per site as units. For ML, the initial tree(s) for heuristic search were also obtained automatically by applying NJ and BioNJ algorithms to pairwise distance matrices estimated using the Maximum Composite Likelihood (MCL) approach, then selecting the topology with superior log likelihood value[61].

Protein structures and preparation

All NMR conformations of *Mex*-LanM (PDB: 6MI5) and calbindin (PDB: 2F33) served as representative replicates. The calbindin structures were truncated to single-out loops EF1–EF2 (residues 16–85) and EF5–EF6 (residues 191–259). The Y^{3+}/Tb^{3+} and Ca^{2+} ions were placed 12~15 Å from the protein surface of the loops under examination. To complete the *Mex*-LanM NMR structures, the amine-terminus alanine residues were added (first residue position) and the carboxyl-terminus polyhistidine-tags were removed. Hydrogen atoms were replaced and both termini capped with an acetyl group (amine-terminus) and an *N*-methyl amide group (carboxyl-terminus). The hydrogen-bond networks were optimized using the Protein Preparation Wizard[65] at pH 7. Several global minimizations by the Maestro software package (Schrödinger 2021-4) were carried out at default settings to remove steric clashes. As the lowest energy structure[11], the first NMR conformation of *Mex*-LanM (PDB: 6MI5) was used as triplicates for MD simulations solely based on EF4-Y^{3+} binding.

MD simulation setup

Prepared structures were solvated in a 12 $Å^3$ orthorhombic box with a TIP3P water model[66,67], neutralized and salted with 0.15 M KCl. The total dimensions approximate 75 × 55 × 65 Å. The Amber-14SB force field[68] was used to parameterize the protein structures. The recently developed force field was applied to parameterize trivalent ions (i.e., Y^{3+}/Tb^{3+})[46]. The previous parameters developed by Li et al. were used for monovalent (i.e., $K^+/Na^+/Cl^-$)[69] and divalent (i.e., Ca^{2+}) ions[70]. The 12-6 Lennard–Jones parameter set from the force fields[46,69,70] conforms with experiments in terms of coordination number for mono-, di- and trivalent ions[47,69]. These ionic parameter sets from Li et al. [46,69,70] were also distinguished to compensate hydration free energy (HFE) while reproducing ion-oxygen distances (IOD). The parameter set under HFE were used for monovalent ions[69] and those under IOD for di- and trivalent ions[46,70].

General MD protocol

All MD simulations were performed using a GPU-accelerated workstation implementing the Desmond software[71]. The parameterized systems were first equilibrated with Desmond default protocols. After equilibration, the final MD stage was conducted under isotropic conditions with an NPT ensemble coupled with a Nose–Hoover thermostat[72] and Martyna-Tobias-Klein barostat[73]. The temperature (300 K) was set with a RESPA[74] integrator at an inner time step of 2-fs. Trajectories were analyzed using the Visual Molecular Dynamics (VMD) software[75] and images captured using VMD or the Maestro software package (Schrödinger 2021-4). The VMD plugin, Hbonds, was used to calculate the first shell water coordination number of Y^{3+} or Ca^{2+} during the MD simulations. The H-bond threshold cutoffs were set at 3.5 Å distance and 180° angle. Results were divided by two to achieve the O_2 coordination number since the Hbonds plugin calculates hydrogen distances.

Ionic competition MD protocol

The Desmond default relaxation protocol was supplemented for in-silico ionic competition with six equilibration steps. Steps 1–5 include an NVT ensemble with restraints on heavy solute atoms (force constant = 30 kcal mol^{-1} Å^{-1}) at a temperature of 10 K for 100-ps each (steps). Step 6 then initiates an annealing simulation at 10 K–300 K for 1000-ps. The last equilibration steps originate from the Desmond default protocol. The final MD production stage was then performed as described in the *General MD protocol*.

LanM homolog sequence identification

A PHI-BLAST[40] using the *Mex*-LanM sequence (PDB: 6MI5) and its EF-hand loop motif, D_1-P_2-D_3-x-D_5-x(3)-D_9-x(2)-E_{12}, was performed against the NCBI non-redundant database that excluded the family of *M. extorquens*, i.e., *Methylobacteriaceae*. Note that the T_7 position was left as a PHI-BLAST[40] wildcard due to its high variance in proteins with EF-hand motifs. Firstly, sequences were removed that did not express a signal peptide, analyzed via the SignalP[76] server, and/or with $100 < x < 160$ residues. Secondly, duplicates were removed that were 90% similar to each other, as interpreted by the CD-hit software[77]. Lastly, sequences that created gaps when aligned, see next section, with *Mex*-LanM were deleted (total = 52 sequence).

Homolog structural prediction

The Phyre2[78] server under 'Intensive mode' was used to predict LanM proteins, PQQ-dependent (alcohol/methanol) dehydrogenase, and TBDT tertiary structures. The Phyre2 performs a multiple sequence alignment built from remote homologs respective of the query[78]. The predicted tertiary structures were prepared, and the hydrogen-bond network was optimized as aforementioned (*Protein structures and preparation*). The Asp/Glu and Lys/Asn residue sidechains for all homolog tertiary structures maintained their respective negative and positive charges. The His residues were set in cationic protonated states. The predicted *Hyphomicrobium* LanM protein structures were refined with 5 × 50-ns runs using the *General MD protocol*, above. The final frame of the final MD refinement for each LanM protein was used as replicates for subsequent MD simulations (i.e., ionic competition). The TBDT refinement protocol is described below.

In-silico TBDT-siderophore binding and system set-up

The default settings in the SwissDock[79] server was used. The siderophores for in-silico binding were automatically uploaded by providing the server with a corresponding ZINC database[80] accession number. The siderophore conformations within the TBDT cavity, i.e., within 14 Å of the FoxA-FOB conformation[53], with favorable binding scores (i.e., negative values) were chosen for subsequent MD simulations. Since the focus was on initial siderophore binding, the amine-terminus, TBDT TonB-boxes were truncated to reduce computational expenditure. Note that the Fe^{3+} bound siderophore was used for MD preparation and simulations (e.g., ferrioxamine, not desferrioxamine). The siderophore structures were prepared from the PDB, i.e., SE8 (PDB: 3MWF) and manually adjusted FOB (PDB: 6I97) for the FOG analog, followed by superimposing them onto the docked structures. The prepared TBDT bound structures were then solvated in a 12 Å^3 orthorhombic box (total dimensions ≈ 85 × 85 × 130 Å) with a TIP3P water model[66,67], neutralized, salted with 0.5 M $MgCl_2$, and embedded in a phosphatidylcholine (POPC) bilayer. The CHARMM36 force field[81] was used to parameterize the TBDT-siderophore bound complexes and POPC membranes. The recently developed force fields from Li et al. [46,69,70] were used to parameterize monovalent, divalent, and trivalent ions. Lastly, the SwissParam server[82] was used to build siderophore topology and force field parameters.

TBDT-membrane system refinement and MD

The POPC embedded, siderophore-bound, TBDT structures were refined by 10 × 5-ns MD simulations. The refinement steps included the Desmond default membrane relaxation protocol, followed by a subsequent *General MD protocol*, and finalized with a 500-ns MD production stage. Residue interactions were inspected using the H-bond contacts plugin from VMD[75]. The distance cutoff was set at 3.5 Å with a 90° angle threshold. The centroid of interacting aromatic residues were calculated and the distances measured using a π-π stacking Tcl script[83].

Hm-LanM vector construction and expression

All chemical reagents regarding *Hm*-LanM expression, purification, and experimentation were purchased from Sigma-Aldrich, Millipore, BioRad, Invitrogen and/or Fisher Scientific/Bioreagent. The gBlocks Gene fragment, encoding the mature protein of *Hm*-LanM (GenBank accession: WP_181335629, 60-136), and the primers for DNA PCR fragments were purchased from Integrated DNA Technologies. The primer set for generating the recombinant protein with Hisx6 tag (*Hm*-LanM_His) was sense 5′- ACT TCC ACC GGA TCC TTA CTC CAC AAG AAG CAG GAA TGA CTT C-3′/antisense 5′- CTC TAG AAT CGA AGG TCG TGC ATC CGC TTC GCT GA-3′. The gBlock Gene fragments were a template for inserting the DNA into Nde-I/BamH-I or Xba-I/BamH-I sites of pT7.JLH plasmid implementing the Gibson assembly master mix (New England BioLab). *E. coli* Rosetta (DE3) cells (Novagen) were transformed with the *Hm*-LanM_His expression vector followed by inoculation onto LB agar plates containing supplement 1 (0.5 mM $CaCl_2$, 0.5 mM $MgCl_2$ and 100 µg/ml ampicillin). A single selected colony was cultured in 5 ml LB broth with supplement 1 at 0.1 mM $CaCl_2$. This starter culture was added into a 500 mL LB broth with supplement 1, then agitated at 200 rpm, and incubated at 37 °C until A600 was 0.8. Protein expression was induced with 0.5 mM IPTG. The *E. coli* cells were pelleted by centrifugation, the supernatant discarded, and the pellet stored at −20 °C until subsequent analyses.

Hm-LanM protein purification

All buffers were prepared using ultrapure water treated with Chelex 100 and incubated at 4 °C. Cell lysis was performed using sonication in 20 mL of *buffer A* (50 mM Tris, 10 mM EDTA, 7.5 mg/mL lysozymes, and 1 mM PMSF, pH 8) and then sonicated using a W-385 heat system. After sonication, insoluble pellets were removed by centrifugation at 20,000 × *g* (30 min). The supernatant was transferred into a snakeskin 3500MWCO tubing and dialyzed against 2 L of *buffer B* (20 mM Tris, 5 mM imidazole, 0.5 M NaCl, pH 8). The dialyzed *E. coli* lysate was pelleted as before and the cleared lysate containing *Hm*-LanM_His was purified using a HisPur Ni-NTA Resin column (1.5 × 2 cm; ThermoFisher) equilibrated with *buffer B*. The protein bound to the column was washed and eluted with *buffer B* at 50 mM Tris. The purified *Hm*-LanM_His was transferred into a snakeskin tubing and dialyzed against 1 L *buffer C* (10 mM EDTA, 20 mM Tris, pH 8) to remove any protein-bound metal traces. The purified *Hm*-LanM_His (10 mg) was then dialyzed against 1 L of Factor Xa (100 µg) digestion in *buffer D* (20 mM Tris, 0.1 M NaCl, 2 mM $CaCl_2$, pH 8).

Hm-LanM purification after Hisx6-tag removal

The mature *Hm*-LanM was filtered through a Ni-NTA column (1.5 × 1 cm), equilibrated with *buffer D* without 2 mM $CaCl_2$, to remove Hisx6-tag and undigested protein. The flow-through fractions were collected and Factor Xa was excised by passing the mature protein through an Amicon Ultra-4 30 kDa MWCO filter. The filtrate was transferred into a snakeskin 3500MWCO tubing and dialyzed against 1 L of *buffer C* to remove leached nickel from Ni-NTA purification. After Ni-removal, the mature *Hm*-LanM was dialyzed against experimentation buffers: *buffer E*, pH 7 (10 mM MOPS, 0.1 M KCl) for CD analysis; or *buffer F* (50 mM Tris, pH 8) for electrophoresis analysis. After the final dialysis, the mature *Hm*-LanM for each experimentation buffer was concentrated with an Amicon Ultra-4 10 kDa MWCO filter.

CD spectral analysis

The spectra analyses of *Hm*-LanM (230 μg/mL) were collected using the Olis DSM 17 Circular Dichroism spectrometer with a 1 mm cell for a 400 μL sample to scan wavelengths 260 to 195 nm. The *Hm*-LanM spectra was collected using *buffer E* in the presence of ~0.05 metal equivalent or 1 μM $LaCl_3$, $YbCl_3$, or $EuCl_3$ per *Hm*-LanM.

Gel electrophoresis experiments

The *Hm*-LanM protein (100 μg/mL) was incubated in *buffer F* with the specified concentrations of REEs or Ca^{2+}. The band mobility shifts were observed by electrophoresis using a 15% acrylamide gel and/or containing 0.05% Sarkosyl and stained with CBB R-250. For MW determination under REEs or Ca^{2+} treatment, the SDS in precision plus protein standards was removed by methanol-chloroform precipitation and reconstituted with Sarkosyl gel loading buffer. Gel images were captured with the ImageQuant 300 (GE Healthcare Biosciences).

Statistics and reproducibility

Formal statistical analyses were not applied given the exploratory nature of the datasets. Visualization techniques such as normalized distributions, boxplots, scatter plots, and line graphs were used to illustrate means, outliers, trends and variations. For hydrochemistry, during environmental DNA (eDNA) sampling and for establishing baseline geochemical profiles across the lake water column, water samples were collected at eight depths (47, 48, 48.5, 49, 50, 52, 54, and 55 m). To assess consistency across measurements, replicates were taken at four depths (47, 48.5, 50, and 54 m). Cation analyses (i.e., dissolved Fe, Mn, and Ln) were conducted using ICP-MS on sample aliquots, with error of the measurements determined using the replicates at 48.5 m. A blank sample, obtained from the fourth rinse of the sampling apparatus, consistently returned values below the detection limit (<D/L) for all elements, validating sample integrity and precision. For MD simulations, NMR structural conformations served as replicates for the indicated time frames. Replicates for predicted structures were generated from the last frame of several MD refinement simulations. For molecular biology, the protein was isolated to assess secondary structures via CD spectra and visualized in gel mobility shift assays.

Data availability

The microbial amplicon dataset of Lake Medard[39] can be found EMBL-EBI under the European Nucleotide Project number, PRJEB47217. The structures for MD analysis, and the sequences for phylogenetic analyses and structure predictions were derived from publicly available databases (i.e., PDB and NCBI). Siderophores for in-silico TBDT binding are available in the ZINC database[80]. The sequence used for *Hm*-LanM vector construction and protein expression for subsequent protein isolation, purification, CD spectra, and gel mobility shift analyses is also publicly available under NCBI GenBank accession number, WP_181335629. Source data provided for *Hm*-LanM CD spectra analysis (Fig. 5A) can be found in Supplementary Data 1. All other data supporting this study are available in the main text and Supplementary Information.

Code availability

The Desmond software[71] under the academic version of the Maestro platform (Schrödinger 2021-4) was used for MD simulations. The Tcl code for π-π stacking centroid calculation of TBDT interacting aromatic residues (https://doi.org/10.5281/zenodo.6408973)[83] and the R-based genoPlotR package (https://genoplotr.r-forge.r-project.org/)[84] for comparative genomics are publicly available. All other software and web-services are detailed in the main text and Supplementary Information.

Received: 30 March 2024; Accepted: 12 November 2024;

References

1. Ganguli, R. & Cook, D. R. Rare earths: a review of the landscape. *MRS Energy Sustain.* **5**, E9 (2018).
2. Ascenzi, P. et al. Rare earth elements (REE) in biology and medicine. *Rend. Lincei Sci. Fis. Nat.* **31**, 821–833 (2020).
3. U.S. Geological Survey. *Mineral Commodity Summaries*; https://doi.org/10.3133/mcs2021 (2021).
4. Dutta, T. et al. Global demand for rare earth resources and strategies for green mining. *Environ. Res.* **150**, 182–190 (2016).
5. Balaram, V. Rare earth elements: a review of applications, occurrence, exploration, analysis, recycling, and environmental impact. *Geosci. Front.* **10**, 1285–1303 (2019).
6. Deblonde, G. J.-P. et al. Impact of a biological chelator, lanmodulin, on minor actinide aqueous speciation and transport in the environment. *Environ. Sci. Technol.* **57**, 20830–20843 (2023).
7. Jung, H., Su, Z., Inaba, Y., West, A. C. & Banta, S. Genetic modification of *Acidithiobacillus Ferrooxidans* for rare-earth element recovery under acidic conditions. *Environ. Sci. Technol.* **57**, 19902–19911 (2023).
8. Ye, Q., Jin, X., Zhu, B., Gao, H. & Wei, N. Lanmodulin-functionalized magnetic nanoparticles as a highly selective biosorbent for recovery of rare earth elements. *Environ. Sci. Technol.* **57**, 4276–4285 (2023).
9. Cotruvo, J. A., Featherston, E. R., Mattocks, J. A., Ho, J. V. & Laremore, T. N. Lanmodulin: a highly selective lanthanide-binding protein from a lanthanide-utilizing bacterium. *J. Am. Chem. Soc.* **140**, 15056–15061 (2018).
10. Carl-Eric, W. et al. Extracellular and intracellular lanthanide accumulation in the methylotrophic *Beijerinckiaceae* bacterium RH AL1. *Appl. Environ. Microbiol.* **87**, e03144–20 (2021).
11. Cook, E. C., Featherston, E. R., Showalter, S. A. & Cotruvo, J. A. Structural basis for rare earth element recognition by *Methylobacterium Extorquens* lanmodulin. *Biochemistry* **58**, 120–125 (2019).
12. Drake, S. K., Lee, K. L. & Falke, J. J. Tuning the equilibrium ion affinity and selectivity of the EF-hand calcium binding motif: substitutions at the gateway position. *Biochemistry* **35**, 6697–6705 (1996).
13. Veenstra, T. D., Gross, M. D., Hunziker, W. & Kumar, R. Identification of metal-binding sites in rat brain calcium-binding protein. *J. Biol. Chem.* **270**, 30353–30358 (1995).
14. Gutenthaler, S. M. et al. Lanmodulin peptides – unravelling the binding of the EF-hand loop sequences stripped from the structural corset. *Inorg. Chem. Front* **9**, 4009–4021 (2022).
15. Mattocks, J. A. et al. Enhanced rare-earth separation with a metal-sensitive lanmodulin dimer. *Nature* **618**, 87–93 (2023).
16. Featherston, E. R. & Cotruvo, J. A. The biochemistry of lanthanide acquisition, trafficking, and utilization. *Biochim. Biophys. Acta Mol. Cell Res.* **1868**, 118864 (2021).
17. Juma, P. O. et al. Siderophore for lanthanide and iron uptake for methylotrophy and plant growth promotion in *Methylobacterium Aquaticum* strain 22A. *Front. Microbiol.* **13**, https://doi.org/10.3389/fmicb.2022.921635 (2022).
18. Sandy, M. & Butler, A. Microbial iron acquisition: marine and terrestrial siderophores. *Chem. Rev.* **109**, 4580–4595 (2009).
19. Zytnick, A. M. et al. Identification and characterization of a small-molecule metallophore involved in lanthanide metabolism. *Proc. Natl Acad. Sci.* **121**, https://doi.org/10.1073/pnas.2322096121 (2024).
20. Deng, Y. W., Ro, S. Y. & Rosenzweig, A. C. Structure and function of the lanthanide-dependent methanol dehydrogenase XoxF from the methanotroph *Methylomicrobium Buryatense* 5GB1C. *J. Biol. Inorg. Chem.* **23**, 1037–1047 (2018).
21. Groom, J. D., Ford, S. M., Pesesky, M. W. & Lidstrom, M. E. A mutagenic screen identifies a TonB-dependent receptor required for the lanthanide metal switch in the type I methanotroph *Methylotuvimicrobium Buryatense* 5GB1C. *J. Bacteriol.* **201**, https://doi.org/10.1128/JB.00120-19 (2019).

22. Ameyama, M., Matsushita, K., Shinagawa, E., Hayashi, M. & Adachi, O. Pyrroloquinoline quinone: excretion by methylotrophs and growth stimulation for microorganisms. *Biofactors* **1**, 51–53 (1988).
23. Lumpe, H. et al. The earlier the better: structural analysis and separation of lanthanides with pyrroloquinoline quinone. *Chem. A Eur. J.* **26**, 10133–10139 (2020).
24. Pol, A. et al. Rare earth metals are essential for methanotrophic life in volcanic mudpots. *Environ. Microbiol* **16**, 255–264 (2014).
25. Shen, Y.-Q. et al. Distribution and properties of the genes encoding the biosynthesis of the bacterial cofactor, pyrroloquinoline quinone. *Biochemistry* **51**, 2265–2275 (2012).
26. Teizi, U., Kazuya, Y., Hisao, K., Akio, Y. & Chieko, I.-Y. Production of pyrroloquinoline quinone by using methanol-utilizing bacteria. *Appl. Environ. Microbiol.* **58**, 3970–3976 (1992).
27. Nealson, K. H. The manganese-oxidizing bacteria. In *The Prokaryotes* 222–231 (Springer New York: New York, NY, 2006).
28. Petrash, D. A. et al. Aqueous system-level processes and prokaryote assemblages in the ferruginous and sulfate-rich bottom waters of a post-mining lake. *Biogeosciences* **19**, 1723–1751 (2022).
29. Umbría-Salinas, K. et al. Redox-driven geochemical partitioning of metal(Loid)s in the iron-rich anoxic sediments of a recently flooded lignite mine pit: Lake Medard, NW Czechia. *J. Hazard. Mater. Adv.* **3**, 100009 (2021).
30. Garrity, G. M., Bell, J. A. & Lilburn, T. Class I. Alphaproteobacteria Class. Nov. in *Bergey's Manual® of Systematic Bacteriology* 1–574 (Springer US: Boston, MA, 2005).
31. *The Prokaryotes*; (eds Rosenberg, E., DeLong, E. F., Lory, S., Stackebrandt, E., Thompson, F.) (Springer Berlin Heidelberg: Berlin, Heidelberg, 2014). https://doi.org/10.1007/978-3-642-30197-1.
32. Keltjens, J. T., Pol, A., Reimann, J. & Op den Camp, H. J. M. PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference. *Appl. Microbiol. Biotechnol.* **98**, 6163–6183 (2014).
33. Krause, S. M. B. et al. Lanthanide-dependent cross-feeding of methane-derived carbon is linked by microbial community interactions. *Proc. Natl Acad. Sci.* **114**, 358–363 (2017).
34. Picone, N. & Op den Camp, H. J. M. Role of rare earth elements in methanol oxidation. *Curr. Opin. Chem. Biol.* **49**, 39–44 (2019).
35. Tanaka, Y., Yoshida, T., Watanabe, K., Izumi, Y. & Mitsunaga, T. Cloning and analysis of methanol oxidation genes in the methylotroph *Hyphomicrobium Methylovorum* GM2. *FEMS Microbiol. Lett.* **154**, 397–401 (1997).
36. Palermo, C. & Dittrich, M. Evidence for the biogenic origin of manganese-enriched layers in lake superior sediments. *Environ. Microbiol. Rep.* **8**, 179–186 (2016).
37. He, S., Barco, R. A., Emerson, D. & Roden, E. E. Comparative genomic analysis of neutrophilic iron(II) oxidizer genomes for candidate genes in extracellular electron transfer. *Front. Microbiol.* **8**. https://doi.org/10.3389/fmicb.2017.01584 (2017).
38. Cyriaque, V. et al. Metal-induced bacterial interactions promote diversity in river-sediment microbiomes. *FEMS Microbiol. Ecol.* **96** https://doi.org/10.1093/femsec/fiaa076 (2020).
39. Planktonic 16S RRNA Data, Bottom Ferruginous/Sulfate-Rich Water Column of Lake Medard, NW Czechia [Data Set]. *European Nucleotide Archive (EMBL-EBI)* http://www.ebi.ac.uk/ena/browser/view/PRJEB47217 (2021).
40. Zhang, Z. et al. Protein sequence similarity searches using patterns as seeds. *Nucleic Acids Res.* **26**, 3986–3990 (1998).
41. Petrash, D. A., Jan, J., Sirová, D., Osafo, N. O. A. & Borovec, J. Iron and nitrogen cycling, bacterioplankton community composition and mineral transformations involving phosphorus stabilisation in the ferruginous hypolimnion of a post-mining lake. *Environ. Sci. Process Impacts* **20**, 1414–1426 (2018).
42. Good, N. M. et al. Lanthanide-dependent alcohol dehydrogenases require an essential aspartate residue for metal coordination and enzymatic function. *J. Biol. Chem.* **295**, 8272–8284 (2020).
43. Chu, F. & Lidstrom, M. E. XoxF acts as the predominant methanol dehydrogenase in the type I methanotroph Methylomicrobium Buryatense. *J. Bacteriol.* **198**, 1317–1325 (2016).
44. Schmitz, R. A. et al. Neodymium as metal cofactor for biological methanol oxidation: structure and kinetics of an XoxF1-type methanol dehydrogenase. *mBio* **12**, https://doi.org/10.1128/mBio.01708-21 (2021).
45. Barona-Gómez, F., Wong, U., Giannakopulos, A. E., Derrick, P. J. & Challis, G. L. Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces Coelicolor* M145. *J. Am. Chem. Soc.* **126**, 16282–16283 (2004).
46. Li, P., Song, L. F. & Merz, K. M. Jr. Parameterization of highly charged metal ions using the 12-6-4 LJ-type nonbonded model in explicit water. *J. Phys. Chem. B* **119**, 883–895 (2015).
47. Li, P. Bridging the 12-6-4 model and the fluctuating charge model. *Front. Chem.* **9**, 721960 (2021).
48. Li, Z., Song, L. F., Li, P. & Merz, K. M. Parametrization of trivalent and tetravalent metal ions for the OPC3, OPC, TIP3P-FB, and TIP4P-FB water models. *J. Chem. Theory Comput.* **17**, 2342–2354 (2021).
49. Kojetin, D. J. et al. Structure, binding interface and hydrophobic transitions of Ca^{2+}-loaded calbindin-D28K. *Nat. Struct. Mol. Biol.* **13**, 641–647 (2006).
50. Åkerfeldt, K. S., Coyne, A. N., Wilk, R. R., Thulin, E. & Linse, S. Ca^{2+}-binding stoichiometry of calbindin D28k as assessed by spectroscopic analyses of synthetic peptide fragments. *Biochemistry* **35**, 3662–3669 (1996).
51. Huang, L. et al. 05SAR-PAGE: separation of protein dimerization and modification using a gel with 0.05% sarkosyl. *Anal. Chim. Acta* **1101**, 193–198 (2020).
52. Hanpaibool, C. et al. Pyrazolones potentiate colistin activity against MCR-1-producing resistant bacteria: computational and microbiological study. *ACS Omega* **8**, 8366–8376 (2023).
53. Josts, I., Veith, K. & Tidow, H. Ternary structure of the outer membrane transporter FoxA with resolved signalling domain provides insights into TonB-mediated siderophore uptake. *eLife* **8**, e48528 (2019).
54. Ohnuki, T., Kozai, N., Sakamoto, F., Suzuki, Y. & Yoshida, T. Biological change of chemical states of actinides and lanthanides-effects of organic acids. *Energy Procedia* **39**, 175–182 (2013).
55. Crocket, K. C. et al. Rare earth element distribution in the NE Atlantic: evidence for benthic sources, longevity of the seawater signal, and biogeochemical cycling. *Front. Mar. Sci.* **5**, https://doi.org/10.3389/fmars.2018.00147 (2018).
56. Catts, J. G. & Langmuir, D. Adsorption of Cu, Pb and Zn by ΔMnO2: applicability of the site binding-surface complexation model. *Appl. Geochem.* **1**, 255–264 (1986).
57. Valero, A., Jan, J. & Petrash, D. A. Anaerobic dissolved As(III) removal from metal-polluted waters by cathode-stabilized Fe(III)-oxyhydroxides. *Environ. Sci.* **9**, 454–466 (2023).
58. Hao, X. et al. Recent advances in exploring the heavy metal(loid) resistant microbiome. *Comput. Struct. Biotechnol. J.* **19**, 94–109 (2021).
59. Tostevin, R. et al. Low-oxygen waters limited habitable space for early animals. *Nat. Commun.* **7**, 12818 (2016).
60. Moore, R. L. The biology of hyphomicrobium and other prosthecate, budding bacteria. *Annu. Rev. Microbiol.* **35**, 567–594 (1981).
61. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
62. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
63. Felsenstein, J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783 (1985).
64. Kimura, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).

65. Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput. Aided Mol. Des.* **27**, 221–234 (2013).
66. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
67. Mahoney, M. W. & Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys.* **112**, 8910–8922 (2000).
68. Maier, J. A. et al. Ff14SB: improving the accuracy of protein side chain and backbone parameters from Ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
69. Li, P., Song, L. F. & Merz, K. M. Systematic parameterization of monovalent ions employing the nonbonded model. *J. Chem. Theory Comput.* **11**, 1645–1657 (2015).
70. Li, P., Roberts, B. P., Chakravorty, D. K. & Merz, K. M. Rational design of particle mesh ewald compatible lennard-jones parameters for +2 metal cations in explicit solvent. *J. Chem. Theory Comput.* **9**, 2733–2748 (2013).
71. Bowers, K. J. et al. Scalable algorithms for molecular dynamics simulations on commodity clusters. In *SC '06: Proc. 2006 ACM/IEEE Conference on Supercomputing*, 43 (ACM/IEEE: Tampa, FL, USA, 2006).
72. Evans, D. J. & Holian, B. L. The Nose–Hoover thermostat. *J. Chem. Phys.* **83**, 4069–4074 (1985).
73. Martyna, G. J., Tobias, D. J. & Klein, M. L. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177–4189 (1994).
74. Tuckerman, M., Berne, B. J. & Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **97**, 1990–2001 (1992).
75. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
76. Almagro Armenteros, J. J. et al. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
77. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).
78. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
79. Grosdidier, A., Zoete, V. & Michielin, O. SwissDock, a protein-small molecule docking web service based on EADock DSS. *Nucleic Acids Res* **39**, W270–W277 (2011).
80. Irwin, J. J. & Shoichet, B. K. ZINC − a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182 (2005).
81. Huang, J. & MacKerell, A. D. Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).
82. Zoete, V., Cuendet, M. A., Grosdidier, A. & Michielin, O. SwissParam: a fast force field generation tool for small organic molecules. *J. Comput. Chem.* **32**, 2359–2368 (2011).
83. Li, J. Gmxtools. https://doi.org/10.5281/zenodo.6408973 (2022).
84. Guy, L., Roat Kultima, J. & Andersson, S. G. E. GenoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **26**, 2334–2335 (2010).
85. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator: Fig. 1. *Genome Res.* **14**, 1188–1190 (2004).

Acknowledgements

We are grateful to the reviewers of *Communications Biology* for their constructive criticism that significantly improved the earliest versions of this manuscript. This study also benefited from consultation and advice of Victor Guallar (Barcelona Supercomputing Center, Spain) and John Elliott (Alberta Diabetes Institute, University of Alberta). JJV thanks the support provided by the P JAC project "Photomachines" (Reg. No CZ.02.01.01/00/22_008/0004624). DAP was partially supported by funding from the Czech Science Foundation to BC-CAS (Grant No. 19-15096Y) and by MEYS CZ - Operational Programme RDE (SoWa Ecosystem Research; project no. CZ.02.1.01/0.0/0.0/16_013/0001782). KK would like to thank the Natural Sciences and Engineering Research Council of Canada to KK (RGPIN-2020-05189) for their continued financial support. The open access publication fee was covered by the Institute of Parasitology (BC-CAS), Algatech, Institute of Soil Biology and Biogeochemistry (BC-CAS), and the "Research Organization Long-term Development Concept of the Czech Geological Survey for the Period 2023–2027". Recombinant protein expression and experimentation were carried out by Dr. Kunimasa Suzuki at the University of Alberta in MolBioCore (Alberta Diabetes Institute) and Analytical & Instrumentation Laboratory (Department of Chemistry).

Author contributions

Design, analyses, interpretation, and write-up: J.J.V. (molecular biology, MD computations, and bioinformatics), and D.P. and K.O.K. (hydrochemistry and geomicrobiology).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-07258-3.

Correspondence and requests for materials should be addressed to James J. Valdés or Daniel A. Petrash.

Peer review information *Communications Biology* thanks Zhen Li and the other, anonymous, reviewer for their contribution to the peer review of this work. Primary Handling Editors: Aylin Bircan and Ophelia Bu. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints