



3D architecture and structural flexibility revealed in the subfamily of large glutamate dehydrogenases by a mycobacterial enzyme

Melisa Lázaro, Roberto Melero, Charlotte Huet, Jorge P López-Alonso, Sandra Delgado, Alexandra Dodu, Eduardo M Bruch, Luciano A Abriata, Pedro M Alzari, Mikel Valle, et al.

► To cite this version:

Melisa Lázaro, Roberto Melero, Charlotte Huet, Jorge P López-Alonso, Sandra Delgado, et al.. 3D architecture and structural flexibility revealed in the subfamily of large glutamate dehydrogenases by a mycobacterial enzyme. *Communications Biology*, 2021, 4 (1), pp.684. <10.1038/s42003-021-02222-x>. <hal-03878750>

HAL Id: hal-03878750

<https://hal.science/hal-03878750v1>

Submitted on 30 Nov 2022






HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

3D architecture and structural flexibility revealed in the subfamily of large glutamate dehydrogenases by a mycobacterial enzyme

Melisa Lázaro¹, Roberto Melero², Charlotte Huet^{3,9}, Jorge P. López-Alonso¹ , Sandra Delgado¹, Alexandra Dodu¹, Eduardo M. Bruch^{4,9}, Luciano A. Abriata^{5,6} , Pedro M. Alzari³ , Mikel Valle¹  [✉] & María-Natalia Lisa^{7,8,3} 

Glutamate dehydrogenases (GDHs) are widespread metabolic enzymes that play key roles in nitrogen homeostasis. Large glutamate dehydrogenases composed of 180 kDa subunits (L-GDH₁₈₀) contain long N- and C-terminal segments flanking the catalytic core. Despite the relevance of L-GDH₁₈₀ in bacterial physiology, the lack of structural data for these enzymes has limited the progress of functional studies. Here we show that the mycobacterial L-GDH₁₈₀ (mL-GDH₁₈₀) adopts a quaternary structure that is radically different from that of related low molecular weight enzymes. Intersubunit contacts in mL-GDH₁₈₀ involve a C-terminal domain that we propose as a new fold and a flexible N-terminal segment comprising ACT-like and PAS-type domains that could act as metabolic sensors for allosteric regulation. These findings uncover unique aspects of the structure-function relationship in the subfamily of L-GDHs.

¹Center for Cooperative Research in Biosciences (CIC bioGUNE), Basque Research and Technology Alliance (BRTA), Bizkaia Technology Park, Building 801 A, Derio, Spain. ²Centro Nacional de Biotecnología, CNB-CSIC, Darwin 3, Madrid, Spain. ³Unité de Microbiologie Structurale, Institut Pasteur, CNRS UMR 3528, Université de Paris, 25 rue du Docteur Roux, Paris, France. ⁴Integrated Structural Biology Department, IGBMC, 1 Rue Laurent Fries, Illkirch, France. ⁵Laboratory for Biomolecular Modeling, School of Life Sciences, École Polytechnique Fédérale de Lausanne and Swiss Institute of Bioinformatics, Lausanne, Switzerland. ⁶Protein Production and Structure Core Facility, School of Life Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. ⁷Instituto de Biología Molecular y Celular de Rosario (IBR, CONICET-UNR), Ocampo y Esmeralda, Rosario, Argentina. ⁸Plataforma de Biología Estructural y Metabólica (PLABEM), Ocampo y Esmeralda, Rosario, Argentina. ⁹Present address: CH, DBV-Technologies, 177-181 Avenue Pierre Brossolette, 92120 Montrouge, France; EMB, Unité de Microbiologie Structurale, Institut Pasteur, CNRS UMR 3528, Université de Paris, 25 rue du Docteur Roux, Paris, France. ✉email: mvalle@cicbiogune.es; lisa@ibr-conicet.gov.ar

Glutamate dehydrogenases (GDHs) are ubiquitous oligomeric enzymes that catalyze the reversible oxidative deamination of L-glutamate to 2-oxoglutarate, at the crossroad between the Krebs cycle and ammonium assimilation. GDHs are grouped into the subfamily of small GDHs composed of subunits of 50 kDa (S-GDH₅₀) and the subfamily of large GDHs (L-GDHs) composed of monomers of 115 kDa (L-GDH₁₁₅) or 180 kDa (L-GDH₁₈₀)¹. L-GDHs, found in lower eukaryotes and prokaryotes, are NAD⁺ dependent enzymes that differ from S-GDH₅₀ by the presence of long N- and C-terminal extensions flanking the catalytic domain¹. The possible role(s) of such terminal segments in oligomerization and/or enzyme regulation has remained largely unknown^{1–9}.

The relevance of L-GDH₁₈₀ in bacterial physiology has been emphasized in previous studies of environmental⁸ and pathogenic species^{10,11}. Among the later, the mycobacterial L-GDH₁₈₀ (mL-GDH₁₈₀) is part of a signal transduction pathway that senses amino acid availability to control metabolism and virulence of *Mycobacterium tuberculosis*^{7,12,13}. This enzyme is essential for the in vitro growth of the tubercle bacillus^{10,11} whereas it is crucial for *Mycobacterium bovis* BCG survival in media containing glutamate as the sole carbon source¹⁴. Moreover, diverse mechanisms have been implicated in the regulation of L-GDH₁₈₀. The catabolism of glutamate by mL-GDH₁₈₀ is inhibited by the regulator GarA^{6,7} when extracellular nitrogen donor amino acids are available¹² whereas the L-GDH₁₈₀ from *Streptomyces clavuligerus*¹ (filo Actinobacteria, which includes mycobacteria) as well as L-GDH₁₈₀ from Proteobacteria^{2,4,5} are directly regulated by amino acids. Despite the key roles of L-GDH₁₈₀ in the redistribution of amino groups within cells, their 3D structure has remained elusive, preventing a deeper understanding of the molecular basis of enzyme function.

Here we report the 3D structure of the mL-GDH₁₈₀ isoform from *Mycobacterium smegmatis*, obtained through an integrative approach that combined single-particle cryo-EM and X-ray protein crystallography data at resolutions between 3.59 and 6.27 Å. Our findings reveal unique characteristics of domain organization and oligomeric assembly in the L-GDHs subfamily, thus allowing to update the annotation of the Pfam family PF05088 that includes the L-GDH₁₈₀, and offer a rationale for the direct regulation of L-GDH₁₈₀ by metabolites. Furthermore, our cryo-EM data uncover fluctuations of the quaternary structure of mL-GDH₁₈₀ that are possibly relevant for the allosteric regulation of the enzyme activity.

Results

The 3D architecture of mL-GDH₁₈₀. As revealed by X-ray protein crystallography and single-particle cryo-EM (Figs. 1 and 2), mL-GDH₁₈₀ assembles into a homotetramer. mL-GDH₁₈₀ monomers are arranged around perpendicular twofold axes that pass through a central cavity in the structure.

The 6.27 Å resolution crystal structure of the selenomethionine (Se-Met) derivative of mL-GDH₁₈₀ (Fig. 1 and Table 1), obtained as illustrated in Supplementary Fig. 1 through an integrative strategy that also included cryo-EM data up to 3.59 Å, revealed that the protein subunits display, to the best of our knowledge, a unique domain organization (Fig. 1a). The N-terminal segment comprises three ACT (Aspartate kinase-Chorismate mutase-TyrA) -like¹⁵ (hereafter ACT*, see below) domains (ACT*1–3), a PAS (Per-Arnt-Sim) -type¹⁶ domain and three helical motifs (HM1–3). Notably, the primary structures of ACT and PAS domains are poorly conserved and, therefore, these modules are often difficult to identify from BLAST searches^{15,16}. The C-terminal region consists of a single helical domain that showed no detectable structural similarity to previously

characterized proteins in Dali¹⁷, ECOD¹⁸, CATH¹⁹, and VAST²⁰ searches and, therefore, constitutes a possible new fold.

The catalytic domains in the mL-GDH₁₈₀ complex were not found to contribute intersubunit contacts (Fig. 1a). Instead, the N- and C-terminal regions of mL-GDH₁₈₀ provide dimer-like interactions between pairs of monomers. Contacts between mL-GDH₁₈₀ subunits engage the ACT*2, ACT*3, and C-terminal domains (Fig. 1b). Most of the residues involved in interfacial hydrogen bonds or salt bridges in mL-GDH₁₈₀ are strictly conserved in the enzyme isoform from *M. tuberculosis* (O53203, 72% sequence identity)⁷, the L-GDH₁₈₀ from *S. clavuligerus* (E2Q5C0, 47% sequence identity)¹ and the L-GDH₁₁₅ from *Nocardia farcinica* (A0A0H5NTF9, 55% sequence identity over non-gap aligned columns). Except for a single amino acid (Arg560), the same group of residues is also conserved in the L-GDH₁₈₀ from *P. aeruginosa* (Q9HZE0, 40% sequence identity)². These observations underscore the functional relevance of the oligomeric assembly found for mL-GDH₁₈₀.

ACT and PAS modules are known to regulate functionally diverse proteins by driving conformational and/or quaternary structural changes^{15,16}. The binding of specific amino acids to ACT–ACT interfaces confers allosteric control to oligomeric enzymes involved in amino acid metabolism¹⁵ whereas PAS modules sense and transduce chemical or physical stimuli to typically dimeric effector domains¹⁶. The ACT* domains of mL-GDH₁₈₀ differ from the archetypal ACT fold in that strand β₁ is located in the position usually occupied by strand β₄, creating an ACT-like ββαββα topology with a β₁β₂β₄β₃ antiparallel sheet (Fig. 1b and Supplementary Fig. 2). Similar variations of the characteristic ACT fold have been described for aspartate kinases and a mammalian tyrosine hydroxylase^{15,21}, including sixteen core residues that are conserved in the ACT*1–3 domains of mL-GDH₁₈₀ (Supplementary Fig. 2). Notably, the interaction between ACT*3 modules in mL-GDH₁₈₀ produces a continuous eight-stranded antiparallel β-sheet with helices on one side (Fig. 1b). A similar side-by-side arrangement of ACT domains generates allosteric amino acid binding sites in 3-phosphoglycerate synthases and aspartate kinases¹⁵. Close to a dimer-like interface, the PAS module in mL-GDH₁₈₀ adopts a typical fold (Fig. 1c), comprising a core five-stranded β-sheet usually involved in signal sensing¹⁶, and displays up to 12% sequence identity with PAS domains in sensor histidine kinases retrieved in Dali¹⁷ searches.

Similarly to S-GDH₅₀, the catalytic core of mL-GDH₁₈₀ consists of subdomains SDI and SDII (Fig. 1d), with the active site located in a groove in-between. Functionally important residues in the catalytic domain of L-GDH₁₈₀ have been previously identified by their conservation in sequence comparisons of diverse GDHs¹. The SDI in mL-GDH₁₈₀ contains most of the residues of the glutamate-binding region whereas the SDII conforms the dinucleotide-binding site.

Intrinsic flexibility and alternate conformers of mL-GDH₁₈₀. Cryo-EM and SAXS data uncovered the intrinsic flexibility of native mL-GDH₁₈₀ (Fig. 2, Supplementary Fig. 3, and Supplementary Table 1). The domains comprised in the 1–500 region of the protein are stabilized by crystal contacts in the crystallographic structure. In contrast, 2D averages for side views of mL-GDH₁₈₀ tetramers revealed a high degree of flexibility at distal ends, where ACT*1–2 and PAS domains reside, and their corresponding densities vanished in 3D cryo-EM maps (Fig. 2a). A 3D-classification of the detected mL-GDH₁₈₀ particles was performed to distinguish alternate conformers of the enzyme. Two mL-GDH₁₈₀ conformers were found, called the open and closed conformations (Fig. 2b), for which the ACT*3 module, the HM3, the catalytic domain and the C-terminal region were defined in

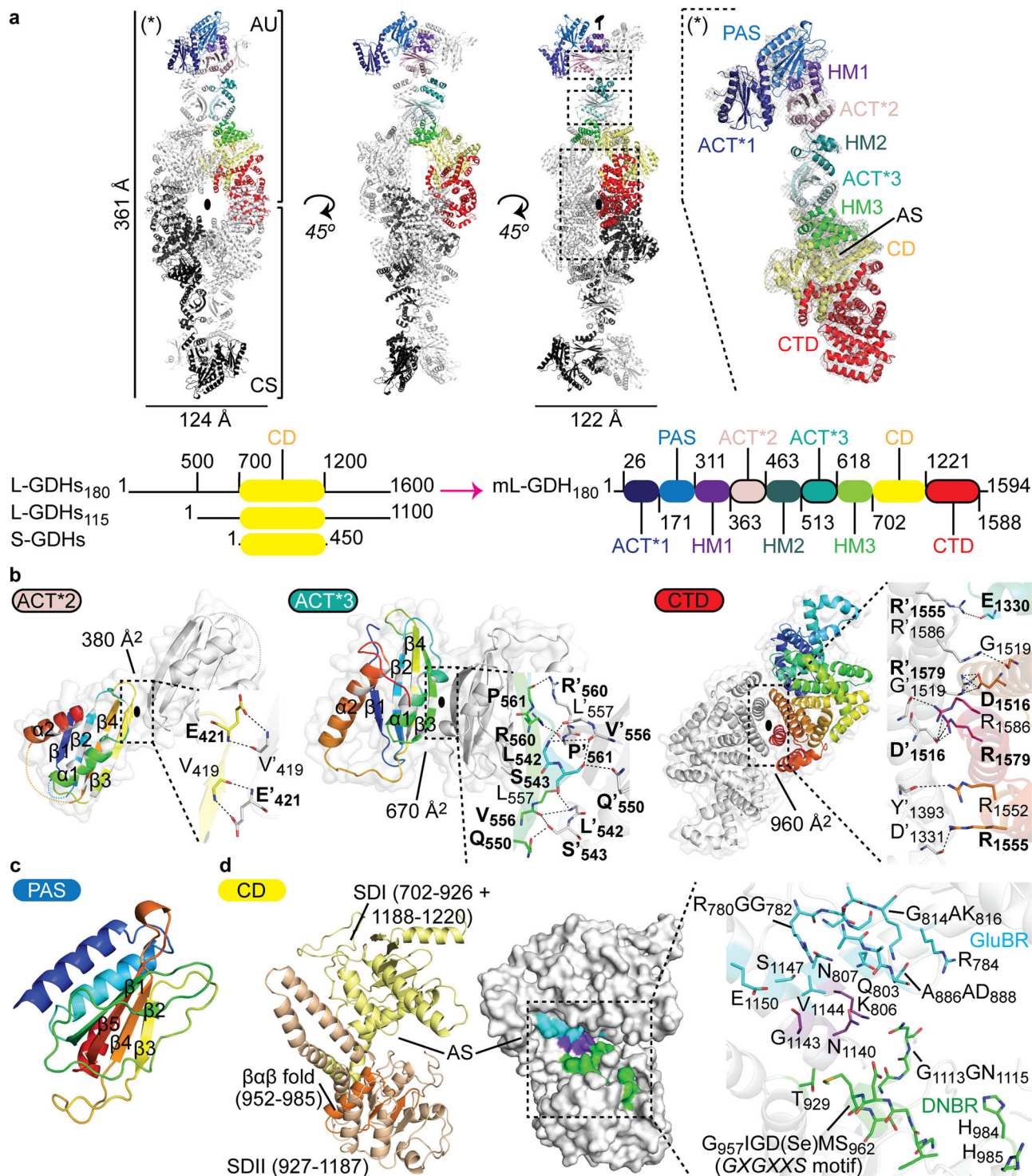


Fig. 1 Crystal structure of Se-Met mLGDH180. **a** The asymmetric unit (AU) contains two monomers (RMSD of 0.51 Å for 289 alpha carbons in segment 45-362, 0.26 Å for 1163 alpha carbons in segment 368-1588); a tetramer (as ribbons) is formed by crystallographic symmetry (CS); oval symbols represent twofold axes. The $2mFo-DFc$ electron density (gray mesh), contoured to 1.5σ , is shown for one protein subunit on the right. Domains boundaries are given in residue numbers in a scheme below; CD catalytic domain, CTD C-terminal domain, AS active site. A comparative scheme of L-GDHs₁₈₀, L-GDHs₁₁₅, and S-GDHs₅₀ is also provided, with approximate residue numbers. **b** Oligomeric interfaces (areas in Å²) involve the domains ACT*2, ACT*3, and CTD. Contacting residues (as sticks in insets) labeled in bold characters are strictly conserved in diverse L-GDHs. The topology of domains ACT*2 and ACT*3 is highlighted with rainbow colors; white positions within the rainbow depict conserved core residues¹⁵. **c** The PAS domain. **d** The CD is shown with the SDI and SDII in yellow and orange, respectively. The $\beta\alpha\beta$ motif is involved in dinucleotide binding¹. The glutamate-binding region (GluBR, cyan) and the dinucleotide-binding region (DNBR, green)¹ are highlighted in a surface representation and as sticks in an inset. Residues in purple conform both binding regions 1. See also Supplementary Figs. 1 and 2.

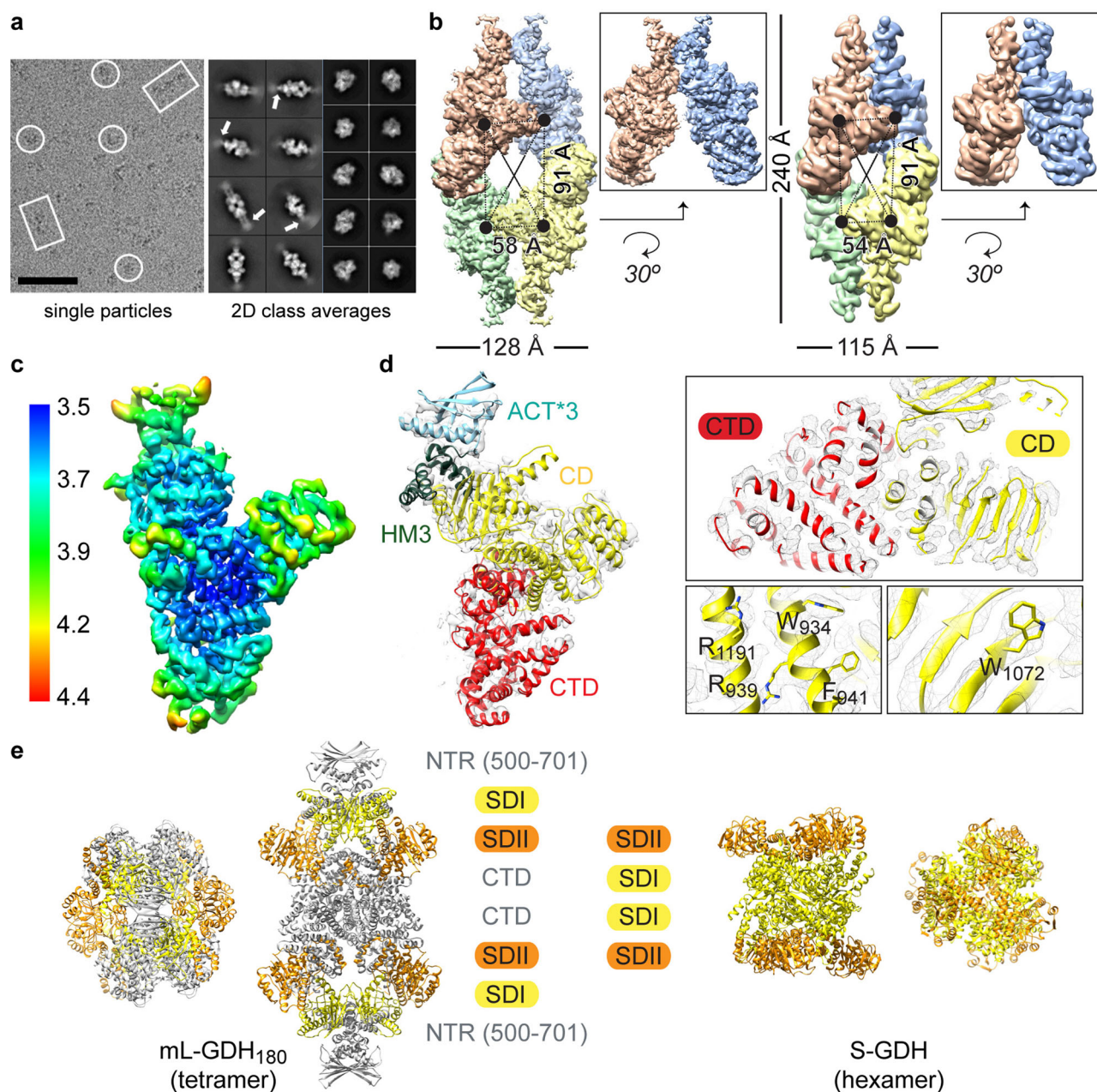


Fig. 2 Intrinsic flexibility and alternate conformers of native mL-GDH180. **a** Cryo-EM image obtained for mL-GDH₁₈₀ showing side (rectangles) and top (circle) views for single particles. Scale bar: 50 nm. The 2D class averages for mL-GDH₁₈₀ tetramers display flexible ends at side views (white arrows). **b** Cryo-EM density maps for the open (left, 4.19 Å resolution) and closed (right, 6.6 Å resolution) conformations of mL-GDH₁₈₀ tetramers, segmented into the four subunits. Black dots: centers of mass of the subunits. Insets are close-up views of the contact zone between the N-terminal regions (NTRs) of two monomers. **c** Local resolution for a single subunit of the open conformation after focused refinement (average resolution is 3.59 Å). **d** Cryo-EM map for one mL-GDH₁₈₀ subunit and the fitted atomic coordinates. Domains colors and labels are as in Fig. 1. Insets are close-up views; selected amino acid side chains are shown as sticks. **e** Comparison of the quaternary structure of mL-GDH₁₈₀, with D2 symmetry, and a representative hexameric S-GDH₅₀ (PDB code 3SBO), with D3 symmetry. The catalytic domains are colored into SDI (yellow) and SDII (orange). The NTRs (only the portion that is well defined in cryo-EM maps is displayed) and the CTDs of mL-GDH₁₈₀ monomers are depicted in gray. See also Supplementary Figs. 3 and 4.

each monomer, achieving an estimated 3.59 Å resolution for this region in the open conformation (Fig. 2c, d, Supplementary Figs. 4 and 5, and Table 2). The two conformers differ in the relative positions of the centers of mass of the subunits (Fig. 2b). The catalytic domains in mL-GDH₁₈₀ monomers in contact through their N-terminal segments are found closer to each other in the less stable closed conformation compared to the open form. Overall, these findings reveal transitions of the quaternary structure that could intervene in the allosteric regulation of the enzyme.

Discussion

L-GDHs₁₈₀ were discovered in 2000 from a study of *Streptomyces clavuligerus*¹ and were later isolated from other diverse bacterial species, including *Pseudomonas aeruginosa*², psychrophilic bacteria^{3–5}, *Caulobacter crescentus*⁸, and *Mycobacterium* spp^{6,7}. As sequences of L-GDHs were identified, they were analyzed in light of the available crystallographic evidence for S-GDHs₅₀^{1,22}.

S-GDHs₅₀ are hexameric enzymes in which the oligomeric interfaces are conformed by motifs that are located within the

Table 1 X-ray diffraction data collection and refinement statistics.

	Se-Met mL-GDH ₁₈₀ (PDB code 7JSR)
Data collection	
Space group	C222 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	151.6, 253.5, 399.7
α , β , γ (°)	90, 90, 90
Resolution (Å)	6.27–24.98 (6.27–7.01) ^a
<i>R</i> _{merge}	0.055 (0.778)
<i>I</i> / σ <i>I</i>	10.4 (1.0)
CC (1/2)	0.999 (0.867)
Completeness (%)	98.4 (100)
Redundancy	4.9 (4.9)
Refinement	
Resolution (Å)	6.27–24.98
No. reflections	16,927
<i>R</i> _{work} / <i>R</i> _{free}	27.7/32.5
No. atoms	
Protein	22,997
<i>B</i> -factors	
Protein	472.38
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.63
Ramachandran	
Favored	93%
Allowed	6.8%
Outliers	0.2%

^aOne protein crystal was employed for structure determination. Values in parentheses are for highest-resolution shell.

catalytic domain^{1,22}. Most of these motifs are substantially modified in L-GDHs, either through sequence changes, insertions, or deletions^{1,22}. In agreement with proposals that the oligomeric assembly would then be different for the two enzyme subfamilies^{1,22}, the quaternary structure of mL-GDH₁₈₀ depends on interactions established by the N- and C-terminal regions flanking the catalytic domain (Fig. 1a) and is radically different from that of S-GDHs₅₀ (Fig. 2e). The stoichiometry of the mL-GDH₁₈₀ complex observed by cryo-EM and X-ray protein crystallography (Figs. 1 and 2) is supported by molecular weight estimates from SAXS data (Supplementary Table 1) and is consistent with previous reports of tetrameric complexes of L-GDHs studied in solution^{2,9}. Furthermore, most of the residues involved in interactions between mL-GDH₁₈₀ monomers are conserved (Fig. 1b) not only in mycobacterial isoforms of the enzyme but also in L-GDHs from diverse species in Actinobacteria and Proteobacteria. Notably, the C-terminal domain, for which we did not find structural similarity with other characterized proteins, has a conserved length among L-GDHs (Fig. 1a) and the contacts between residues predicted from sequence alignments by Raptor X (Supplementary Fig. 6) further support the interactions observed experimentally. All these findings suggest that the oligomeric assembly of mL-GDH₁₈₀ may be a common theme in the enzyme subfamily.

The catalytic domains in the mL-GDH₁₈₀ complex are oriented opposite to those in S-GDHs₅₀ (Fig. 2e), with the SDI (Fig. 1d) directed toward the distal ends of the protein, where the monomers N-terminal region resides. This segment comprises ACT-like modules as well as a PAS-like domain arranged in tandem (Fig. 1a) and shows a high degree of flexibility (Fig. 2a). A comparison of the mL-GDH₁₈₀ conformers identified by cryo-EM (Fig. 2b) shows that conformational changes in the N-terminal region correlate with alterations in the relative positions of the catalytic domains. Taking into account the known roles of ACT modules in the allosteric control of oligomeric enzymes involved in amino acid metabolism¹⁵, our findings offer a rationale for previous evidence pointing out the direct regulation of diverse L-GDHs₁₈₀ by metabolites^{1,2}.

Table 2 Cryo-EM data collection and processing.

	Open form (EMD-11606)	Closed form (EMD-11612)	Monomer(EMD-11613)
Data collection and processing			
Magnification	47,170	47,170	47,170
Voltage (kV)	300	300	300
Electron dose (e ⁻ /Å ²)	40 (20 fractions)	40 (20 fractions)	40 (20 fractions)
Defocus range (μm)	0.67–3.26	0.67–3.26	0.67–3.26
Pixel size (Å)	1.06	1.06	1.06
Symmetry imposed	D2	D2	C1
Initial particles	276,704	276,704	276,704
images (no.)			
Final particles	63,715	42,476	63,715 × 4
images (no.)			
Map resolution (Å)	4.19	6.6	3.59
FSC threshold	0.143	0.143	0.143
Map sharpening	−118.9	−250	−118.9
<i>B</i> factor			
Refinement			
Initial model used	7JSR	–	–
(PDB code)			
Model resolution (Å)	4.03	–	–
FSC threshold	0.143	–	–
Map sharpening <i>B</i> factor (Å ²)	−118.9	–	–
Model composition			
Non-hydrogen atoms	33,032	–	–
Protein residues	4,252	–	–
Ligands	0	–	–
<i>B</i> factors (Å ²)			
Protein	116.28	–	–
Ligand	–	–	–
R.m.s. deviations			
Bond lengths (Å)	0.006	–	–
Bond angles (°)	1.026	–	–
Validation			
MolProbity score	1.93	–	–
Clashscore	7.87	–	–
Poor rotamers (%)	0.35	–	–
Ramachandran plot			
Favored (%)	91.6	–	–
Allowed (%)	8.4	–	–
Disallowed (%)	0	–	–

In conclusion, our findings suggest that the N-terminal segment of mL-GDH₁₈₀ (as well as in related enzymes) could transduce intracellular metabolic stimuli to the catalytic core by driving changes in the quaternary structure. The reported 3D model of mL-GDH₁₈₀ can now frame future studies to dissect the structure-function relationship of this enzyme and other members of the L-GDHs subfamily.

Methods

Protein production and purification. The sequence coding for the L-GDH₁₈₀ from *M. smegmatis* MC²-155 (MSMEG_4699, Uniprot A0R1C2) was cloned into vector pLIC-His²³ employing the oligonucleotides Fw: CCAGGGAGCAGCCTC GATGATTCGCGGCTTTCGG and Rv: GCAAAGCACCGGCTCGTTACCC AGTCGTTCCGGTCCC. The resulting plasmid was used to produce N-terminally His6-tagged mL-GDH₁₈₀ in *E. coli* cells. Transformed *E. coli* cells were grown at 37 °C in medium supplemented with ampicillin or carbenicillin until reaching 0.8 units of optical density at 600 nm. Protein expression was then induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM, and the incubation was continued for 18 h at 14 °C. Cells were harvested by centrifugation and sonicated. Following clarification by centrifugation, the supernatant was loaded onto a HisTrap HP column (GE Healthcare) equilibrated with buffer 25 mM HEPES, 500 mM NaCl, 20% v/v glycerol, 20 mM imidazole, pH 8.0, and His6-tagged mL-GDH₁₈₀ was purified by applying a linear imidazole gradient (20–500 mM). The protein was then further purified by size-exclusion chromatography, as described below. mL-GDH₁₈₀ containing fractions, as confirmed by SDS-PAGE and measurements of glutamate dehydrogenase activity⁶, were pooled and used immediately. The protein was quantified by electronic absorption using the molar absorption coefficient of 171,090 M^{−1} cm^{−1}, predicted from the amino acid sequence by the ProtParam tool (<http://web.expasy.org/protparam/>).

For EM and SAXS experiments, native mL-GDH₁₈₀ was produced in *E. coli* BL21(DE3) cells grown in LB broth. Size-exclusion chromatography was performed using a Superose 6 10/300 GL column (GE Healthcare) equilibrated in buffer 20 mM MES, 300 mM NaCl, 5 mM MgCl₂, pH 6.0. Instead, Se-Met mL-GDH₁₈₀ for crystallographic studies was produced in *E. coli* B834 (DE3) cells grown in SelenoMethionine Medium Complete (Molecular Dimensions), and size-

exclusion chromatography was carried out using a HiPrep Sephacryl S-400 HR column (GE Healthcare) equilibrated in buffer 25 mM Tris, 150 mM NaCl, pH 7.5.

GarA from *M. tuberculosis* was produced as described by England et al.²⁴.

Cryo-electron microscopy. In all, 4 μ l of 0.3 mg/ml mL-GDH₁₈₀ were applied to Quantifoil R2/2 holey carbon grids and vitrified using a Vitrobot (FEI). Data collection was carried out in a Titan Krios FEI electron microscope operated at 300 kV by a K2 direct detector (GATAN) (eBIC, Electron Bio-Imaging Centre, Diamond light source, Oxford). Movie frames (1,802) were taken at a nominal magnification of $\times 47,170$ resulting in a sampling of 1.06 Å/pixel. Each movie contained 20 frames with an accumulated dose of 40 e⁻/Å². Movie frames were aligned using MotionCor2²⁵, and the final average included frames 2–15 with a total dose of 28 e⁻/Å² on the sample.

The contrast transfer function (CTF) of the micrographs was estimated using CTFIND4²⁶. The particles were automatically selected from the micrographs using autopicking from RELION-3²⁷. Evaluation of the quality of particles and selection was performed after 2D classifications with SCIPION²⁸ and RELION-3²⁷ software packages. The initial volume for 3D image processing was calculated using common lines in EMAN²⁹ and using the algorithm 3D-RANSAC³⁰. With this initial reference, additional rounds of automated particle picking were performed. An initial data set of 276,704 particles was subjected to 2D and 3D class averaging in order to select the best particles. The 3D-classification of the 106,190 final particles with imposed D2 symmetry resulted in two different conformations, a closed (40%) and an open form (60%), with estimated resolutions of 6.6 Å and 4.47 Å, respectively. The set of particles for the open tetramer was further refined after particle polishing in RELION-3^{27,31} over dose-weighted frames (total set of 20 frames), resulting in a 3D EM map at 4.19 Å. A focused refinement on the core of the subunits (excluding blurred regions at the tip ends) further improved the resolution to 3.59 Å for a monomer in the open conformation. This refinement focused on single mL-GDH₁₈₀ subunits was performed after the alignment of all the monomers following the D2 symmetry, with masked subunits. Local resolution was estimated using RELION-3^{27,31}.

Model fitting into cryo-EM maps was performed using the programs UCSF Chimera³², Namdinator³³, phenix.real_space_refine³⁴ and Coot³⁵. Residues 500–1588 from the crystal structure of Se-Met mL-GDH₁₈₀ (see below) were fitted into the cryo-EM map of the open form of the protein. Se-methionine residues were replaced by methionine residues using Coot³⁵ and the model was finally refined employing phenix.real_space_refine³⁴ with NCS and secondary structure restraints. Overall correlation coefficients were: CC (mask): 0.73; CC (box): 0.85; CC (volume): 0.73; and CC (peaks): 0.67. The final model contained 91.6% of the residues within favored regions of the Ramachandran plot and no outliers.

Figures were generated and rendered with UCSF Chimera³².

Cryo-EM maps obtained for mL-GDH₁₈₀ were deposited in the Electron Microscopy Data Bank under the accession codes EMD-11606 (open conformation), EMD-11612 (closed conformation), and EMD-11613 (monomer). Atomic coordinates for the open form of mL-GDH₁₈₀ derived from cryo-EM data were deposited in the Protein Data Bank under the accession code 7A1D.

Negative staining electron microscopy. Negative-stained grids of mL-GDH₁₈₀ were prepared using 2% uranyl acetate and visualized on a JEM-1230 transmission electron microscope (JEOL Europe) at an acceleration voltage of 80 kV. Images were taken in low dose conditions at a nominal magnification of $\times 30,000$ using a GATAN CCD camera, resulting in 2.3 Å/pixel sampling.

Labeling of N-terminally His6-tagged mL-GDH₁₈₀ was performed by direct incubation of electron microscopy grids in solutions containing 5 nm Ni-NTA-Nanogold (Nanoprobes). Briefly, after glow discharging the grids, the protein was incubated for 1 min on the grids, fixed with 2% paraformaldehyde for 10 min at 4 °C, washed 5 min with PBS, incubated for 15 min with Nanogold diluted 1/75 in PBS, washed twice with PBS, and finally stained with 2% uranyl acetate for 45 s.

Crystallization, X-ray data collection, and structure determination. Crystallization screenings were carried out using the sitting-drop vapor diffusion method and a Mosquito (TTP Labtech) nanoliter-dispensing crystallization robot. Crystals of Se-Met mL-GDH₁₈₀ grew after 4–6 months from a 16.5 mg/ml protein solution containing an equimolar amount of GarA from *M. tuberculosis*, by mixing equal volumes of protein solution and mother liquor (100 mM sodium cacodylate pH 5.8, 12% v/v glycerol, and 1.25 M (NH₄)₂SO₄), at 4 °C. Single crystals were cryoprotected in mother liquor containing 32% v/v glycerol and flash-frozen in liquid nitrogen. X-ray diffraction data were collected at the synchrotron beamline ID23-1 (European Synchrotron Radiation Facility, Grenoble, France), at 100 K, using wavelength 0.99187 Å. Diffraction data were processed using XDS³⁶ and scaled with Aimless³⁷ from the CCP4 program suite³⁸.

The crystal structure of Se-Met mL-GDH₁₈₀ was solved by molecular replacement using the program Phaser³⁹. As search probe we used the atomic coordinates of a model built as follows. First, a poly-Ala model of mL-GDH₁₈₀ was obtained from a preliminary ca. 7 Å resolution cryo-EM map of the protein, by employing the program phenix.map_to_model⁴⁰. Features of the catalytic domain in mL-GDH₁₈₀ monomers became apparent in the model, suggesting that the N-terminus of the polypeptide chains was located at the tips of the particle. This was

confirmed by labeling N-terminally His6-tagged mL-GDH₁₈₀ with Ni-NTA-Nanogold (Nanoprobes) and visualizing particles by negative staining electron microscopy. Then, the catalytic domain of mL-GDH₁₈₀ (residues 702–1220) was homology-modeled by using the structure of the S-GDH₅₀ from *C. glutamicum* (PDB code 5GUD) as template and employing MODELLER⁴¹ as implemented in the HHpred server⁴². One copy of the model of the catalytic domain was rigid-body fitted into the 7 Å cryo-EM map of mL-GDH₁₈₀, which allowed updating the starting poly-Ala model by correcting helical elements and incorporating strands corresponding to the catalytic domain in one monomer of mL-GDH₁₈₀. From this, the D2 tetramer was then rebuilt by applying NCS operators detected by phenix.find_ncs⁴³ and the model was refined against the 7 Å cryo-EM map using phenix.real_space_refine³⁴ with NCS and secondary structure restraints. Finally, one of the protein chains in the resulting model was used as search probe to solve the crystal structure of Se-Met mL-GDH₁₈₀ by molecular replacement.

Two monomers were placed within the asymmetric unit, which taken together with nearby crystallographic symmetry mates replicate the quaternary structure observed by cryo-EM. After crystallographic refinement using phenix.refine^{44,45} with NCS and secondary structure restraints, mFo-DFc and 2mFo-DFc electron density maps displayed rod-shaped electron density peaks that remained unmodeled at this stage and that most likely corresponded to helices in the N-terminal region of mL-GDH₁₈₀. Phase improvement by density modification with RESOLVE⁴⁶ provided additional evidence in support of such elements. The N-terminal segment of mL-GDH₁₈₀ (residues 1–701) was modeled ab initio using RaptorX^{47,48}, one of the top-ranking ab initio structure prediction methods according to recent CASP evaluations^{49,50}. RaptorX works by initially estimating residue-residue contacts from residue coevolution patterns and uses the predicted contacts to drive model building; such technique has proven highly successful especially when integrated with experimental data (multiple examples overviewed in Abriata et al.⁵¹). The residue–residue contact map predicted by RaptorX and the models produced from it revealed that the N-terminal segment of mL-GDH₁₈₀ comprises an array of contiguous domains, which were subsequently individually rigid-body fitted into the electron density maps. Similarly, the C-terminal domain of mL-GDH₁₈₀ (residues 1221–1594) was modeled ab initio employing RaptorX^{47,48} and used to correct and complete the crystallographic model. Finally, un-modeled or poorly modeled segments in the CD were manually built employing Coot³⁵ from a 3.59 Å resolution cryo-EM map obtained for a monomer of mL-GDH₁₈₀. The structure was then further refined by iterative cycles of manual model building with Coot³⁵, used to apply stereochemical restraints, and crystallographic refinement of atomic coordinates and individual B-factors using phenix.refine^{44,45} with NCS and secondary structure restraints. The final model contained 93% of the residues within favored regions of the Ramachandran plot and 0.2% of outliers. The crystallographic structure of Se-Met mL-GDH₁₈₀ correctly explained the connecting loops and bulky amino acid side chains evidenced for residues 500–1588 by a 4.19-Å cryo-EM map of the protein, which allowed to validate the strategy used for model building. Furthermore, the position of Se-Met residues in the crystal structure of Se-Met mL-GDH₁₈₀ matched the position of peaks in an anomalous difference map calculated with diffraction data acquired at 0.979338 Å (12.66 keV), the Se K-edge.

Even though Se-Met mL-GDH₁₈₀ crystallized in the presence of GarA from *M. tuberculosis*, electron density maps did not reveal evidences of co-crystallization and molecular replacement attempts with Phaser³⁹ using the atomic coordinates of GarA in PDBs 2KFU or 6I2P failed. The evidence of helical elements in all mL-GDH₁₈₀ domains allows excluding the presence of GarA (an all beta protein) from modeled regions, particularly from those involved in crystal contacts (mL-GDH₁₈₀ residues 1–500). The crystallization of a protein from a mixture of two or more proteins is not an unusual phenomenon, and it has even been reported that a protein can crystallize in a different space group due to the presence of other proteins in the sample without giving rise to co-crystals⁵², just to mention one example.

Figures were generated and rendered with UCSF Chimera³² or Pymol version 1.8.x (Schrödinger, LLC).

Atomic coordinates and structure factors obtained for Se-Met mL-GDH₁₈₀ were deposited in the Protein Data Bank under the accession code 7JSR.

Small angle X-ray scattering. Synchrotron SAXS data were collected at BioSAXS ID14EH3 beamline (European Synchrotron Radiation Facility, Grenoble, France) and recorded at 15 °C using a PILATUS 1 M pixel detector (DECTRIS) at a sample-detector distance of 2.43 m and a wavelength of 0.931 Å, resulting momentum transfer (*s*) ranging from 0.009 to 0.6 Å⁻¹.

mL-GDH₁₈₀ was assayed at concentrations ranging from 1 to 14 mg/ml in buffer 25 mM Tris, 150 mM NaCl, pH 7.5. For the buffer and the samples, 10 2D scattering images were acquired and processed to obtain radially averaged 1D curves of normalized intensity versus scattering angle. In order to optimize background subtraction, buffer scattering profiles recorded before and after measuring every sample were averaged. Then, for each protein sample, the contribution of the buffer was subtracted. All subsequent data processing was performed using the ATSAS suite⁵³.

Average scattering curves corresponding to different protein concentrations were compared using PRIMUS^{53,54}. To obtain the idealized scattering curve the low *s* region of the most diluted sample and the high *s* region of the most

concentrated sample were merged. The values of the forward scattering intensity $I(0)$, the radius of gyration R_g , as well as the dimensionless Kratky plot were calculated using PRIMUS^{53,54}. Guinier plots of independent average scattering curves evidenced a constant R_g at different protein concentrations. The Porod volume was estimated using DATPOROD⁵³ and an s_{\max} value equal to $7.5/R_g$. The pairwise distance distribution function $p(r)$ and the maximum particle dimension D_{\max} were calculated using GNOM^{53,55} with a reduced χ^2 value of 1.07 for curve fitting. After running DAMMIN^{53,56} the excluded volume was estimated as $V_{\text{ex}} = \text{volume of a single dummy atom} \times \text{number of dummy atoms} / 0.74$. Finally, the MW was estimated from the Porod volume and the excluded volume.

Statistics and reproducibility. One protein crystal was employed for structure determination. For EM and SAXS studies, a protein batch was prepared in each case immediately before the experiment. No data were excluded from the analyses. All attempts of replication were successful.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Cryo-EM maps obtained for mL-GDH₁₈₀ were deposited in the Electron Microscopy Data Bank under the accession codes EMD-11606 (open conformation), EMD-11612 (closed conformation), and EMD-11613 (monomer). Atomic coordinates for the open form of mL-GDH₁₈₀ derived from cryo-EM data were deposited in the Protein Data Bank under the accession code 7A1D. Atomic coordinates and structure factors obtained for Se-Met mL-GDH₁₈₀ by X-ray protein crystallography were deposited in the Protein Data Bank under the accession code 7JSR.

Received: 13 December 2020; Accepted: 14 May 2021;

Published online: 03 June 2021

References

- Miñambres, B., Olivera, E. R., Jensen, R. A. & Luengo, J. M. A new class of glutamate dehydrogenases (GDH): biochemical and genetic characterization of the first member, the AMP-requiring NAD-specific GDH of *Streptomyces clavuligerus*. *J. Biol. Chem.* **275**, 39529–39542 (2000).
- Lu, C. D. & Abdelal, A. T. The *gdhB* gene of *Pseudomonas aeruginosa* encodes an arginine-inducible NAD⁺-dependent glutamate dehydrogenase which is subject to allosteric regulation. *J. Bacteriol.* **183**, 490–499 (2001).
- Camardella, L. et al. The Antarctic *Psychrobacter* sp. TAD1 has two cold-active glutamate dehydrogenases with different cofactor specificities. Characterisation of the NAD⁺-dependent enzyme. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **131**, 559–567 (2002).
- Kawakami, R., Sakuraba, H. & Ohshima, T. Gene cloning and characterization of the very large NAD-dependent L-glutamate dehydrogenase from the psychrophile *Janthinobacterium lividum*, isolated from cold soil. *J. Bacteriol.* **189**, 5626–5633 (2007).
- Kawakami, R., Oyama, M., Sakubara, H. & Ohshima, T. The unique kinetic behavior of the very large NAD-dependent glutamate dehydrogenase from *Janthinobacterium lividum*. *Biosci. Biotechnol. Biochem.* **74**, 884–887 (2010).
- O'Hare, H. M. et al. Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol. Microbiol.* **70**, 1408–1423 (2008).
- Nott, T. J. et al. An intramolecular switch regulates phosphoindependent FHA domain interactions in mycobacterium tuberculosis. *Sci. Signal.* **2**, ra12 (2009).
- Beaufay, F. et al. A NAD-dependent glutamate dehydrogenase coordinates metabolism with cell division in *Caulobacter crescentus*. *EMBO J.* **34**, 1786–1800 (2015).
- Veronese, F. M., Nyc, J. F., Brown, D. M. & Smith, E. L. Nicotinamide adenine dinucleotide-specific glutamate dehydrogenase of *Neurospora*. *J. Biol. Chem.* **249**, 7922–7929 (1974).
- Griffin, J. E. et al. High-resolution phenotypic profiling defines genes essential for mycobacterial growth and cholesterol catabolism. *PLoS Pathog.* **7**, e1002251 (2011).
- deJesús, M. A. et al. Bayesian analysis of gene essentiality based on sequencing of transposon insertion libraries. *Bioinformatics* **29**, 695–703 (2013).
- Rieck, B. et al. PknG senses amino acid availability to control metabolism and virulence of *Mycobacterium tuberculosis*. *PLoS Pathog.* **13**, e1006399 (2017).
- York, A. Bacterial physiology: an inside job on metabolism. *Nat. Rev. Microbiol.* **15**, 383–383 (2017).
- Gallant, J. L., Viljoen, A. J., Van Helden, P. D. & Wiid, I. J. F. Glutamate dehydrogenase is required by *Mycobacterium bovis* BCG for resistance to cellular stress. *PLoS ONE* **11**, 1–14 (2016).
- Lang, E. J. M., Cross, P. J., Mittelstädt, G., Jameson, G. B. & Parker, E. J. Allosteric ACTin: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr. Opin. Struct. Biol.* **29**, 102–111 (2014).
- Möglich, A., Ayers, R. A. & Moffat, K. Structure and signaling mechanism of Per-ARNT-sim domains. *Structure* **17**, 1282–1294 (2009).
- Holm, L. DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140 (2020).
- Cheng, H. et al. ECOD: an evolutionary classification of protein domains. *PLoS Comput. Biol.* **10**, e1003926 (2014).
- Dawson, N. L. et al. CATH: an expanded resource to predict protein function through structure and sequence. *Nucleic Acids Res.* **45**, D289–D295 (2017).
- Madej, T. et al. MMDB and VAST+: Tracking structural similarities between macromolecular complexes. *Nucleic Acids Res.* **42**, D297–D303 (2014).
- Zhang, S., Huang, T., Ilangoan, U., Hinck, A. P. & Fitzpatrick, P. F. The solution structure of the regulatory domain of tyrosine hydroxylase. *J. Mol. Biol.* **426**, 1483–1497 (2014).
- Britton, K. L., Baker, P. J., Rice, D. W. & Stillman, T. J. Structural relationship between the hexameric and tetrameric family of glutamate dehydrogenases. *Eur. J. Biochem.* **209**, 851–859 (1992).
- Cabrera, L. D., Dai, W. & Bottomley, S. P. A family of E. coli expression vectors for laboratory scale and high throughput soluble protein production. *BMC Biotechnol.* **6**, 12 (2006).
- England, P. et al. The FHA-containing protein GarA acts as a phosphorylation-dependent molecular switch in mycobacterial signaling. *FEBS Lett.* **583**, 301–307 (2009).
- Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
- Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
- Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
- de la Rosa-Trevín, J. M. et al. Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol.* **195**, 93–99 (2016).
- Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
- Vargas, J., Álvarez-Cabrera, A. L., Marabini, R., Carazo, J. M. & Sorzano, C. O. S. Efficient initial volume determination from electron microscopy images of single particles. *Bioinformatics* **30**, 2891–2898 (2014).
- Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
- Petersen, E. F. et al. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Kidmose, R. T. et al. Namdinator - Automatic molecular dynamics flexible fitting of structural models into cryo-EM and crystallography experimental maps. *IUCr* **6**, 526–531 (2019).
- Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 531–544 (2018).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
- Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **66**, 125–132 (2010).
- Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
- Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D* **67**, 235–242 (2011).
- McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Terwilliger, T. C., Adams, P. D., Afonine, P. V. & Sobolev, O. V. A fully automatic method yielding initial models from high-resolution cryo-electron microscopy maps. *Nat. Methods* **15**, 905–908 (2018).
- Sali, A. & Blundell, T. Comparative modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1994).
- Zimmermann, L. et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
- Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D Struct. Biol.* **75**, 861–877 (2019).
- Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 352–367 (2012).
- Headd, J. J. et al. Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68**, 381–390 (2012).
- Terwilliger, T. C. et al. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **64**, 61–69 (2007).
- Wang, S., Sun, S., Li, Z., Zhang, R. & Xu, J. Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLoS Comput. Biol.* **13**, e1005324 (2017).
- Xu, J. Distance-based protein folding powered by deep learning. *Proc. Natl Acad. Sci. USA* **116**, 16856–16865 (2019).

49. Abriata, L. A., Tamò, G. E., Monastyrskyy, B., Kryshtafovych, A. & Dal Peraro, M. Assessment of hard target modeling in CASP12 reveals an emerging role of alignment-based contact prediction methods. *Proteins Struct. Funct. Bioinform.* **86**, 97–112 (2018).
50. Abriata, L. A., Tamò, G. E. & Dal Peraro, M. A further leap of improvement in tertiary structure prediction in CASP13 prompts new routes for future assessments. *Proteins Struct. Funct. Bioinform.* **87**, 1100–1112 (2019).
51. Abriata, L. A. & Dal Peraro, M. State-of-the-art web services for de novo protein structure prediction. *Brief. Bioinform.* **22**, bbaa139 (2020).
52. Simpkin, A. J. et al. SIMBAD: a sequence-independent molecular-replacement pipeline. *Acta Crystallogr. Sect. D* **74**, 595–605 (2018).
53. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).
54. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).
55. Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
56. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886 (1999).

Acknowledgements

The authors acknowledge the support and the use of resources of Instruct, a Landmark ESFRI project and Instruct Image Processing Center (I2PC) through Instruct Access Project PID6258, NVIDIA Corporation for the donation of the Quadro GP100 GPU used for this research, the eBIC Electron Bio-Imaging Centre (Diamond light source, Oxford) for the collection of cryo-EM images and the European Synchrotron Radiation Facility (Grenoble, France) for the collection of diffraction and SAXS data. We thank David Gil, Ariel Mechaly, Ahmed Haouz, and Pascal Le Normand for technical advice and insightful discussions. This work was supported by the grant PICT 2017-1932, from the Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación (Agencia I + D + i, Argentina), received by M.N.L., and the grant PGC2018-098996-B-I00 from the Spanish Ministerio de Ciencia e Innovación, received by M.V.; M.V. thanks the AEI (Agencia Estatal de Investigación) for the Severo Ochoa Excellence Accreditation (SEV-2016-0644).

Author contributions

M.L. produced protein, prepared EM samples, performed EM experiments, processed, and analyzed EM data, and performed structural analyses; R.M. determined the initial

cryo-EM model; C.H. optimized protein production; J.P.L.A. processed and analyzed EM data; S.D. contributed to protein production and the preparation of EM samples; A.D. contributed to protein production; E.M.B. analyzed SAXS data; L.A.A. contributed to ab initio modeling and structural analyses; P.M.A. designed experiments and analyzed data; M.V. designed experiments, attended EM data collection and processing and analyzed results; M.N.L. cloned the gene of mL-GDH₁₈₀, optimized protein production, produced protein, obtained protein crystals, solved the crystal structure of the protein, refined the structure of mL-GDH₁₈₀ obtained by cryo-EM, designed experiments, acquired data, analyzed data, and wrote the paper. All authors read and corrected the paper.

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/s42003-021-02222-x>.

Correspondence and requests for materials should be addressed to M.V. or M.-N.L.

Peer review information *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021