Differences in the regulation mechanisms of the glutamine synthetase from methanogenic archaea unveiled by structural investigations.

Marie-Caroline Müller[1], Olivier N. Lemaire[1], Julia M. Kurth[2,3], Cornelia U. Welte[2], Tristan Wagner[1]*

[1] Microbial Metabolism Research Group, Max Planck Institute for Marine Microbiology, Celsiusstraße 1, 28359 Bremen, Germany.

[2] Department of Microbiology, Institute for Water and Wetland Research, Radboud University, Heyendaalseweg 135, 6525 AJ Nijmegen, The Netherlands.

[3] Microcosm Earth Centre, Philipps-University Marburg and Max Planck Institute for Terrestrial Microbiology, Hans-Meerwein-Str. 4, 35032 Marburg, Germany.

Corresponding author. Email: twagner@mpi-bremen.de

Keywords: Nitrogen-assimilation, allosteric regulation, methanogenic archaea, structural biology, enzyme mechanism.

Abstract

Glutamine synthetases catalyze the ATP-dependent ammonium assimilation, the initial step of nitrogen acquisition that must be tightly regulated to fit cellular needs. While their catalytic mechanisms and regulation are well-characterized in bacteria and eukaryotes, only limited knowledge exists about the archaeal representatives. Here, we natively purified the glutamine synthetases type I-α from *Methanothermococcus thermolithotrophicus* and *Methermicoccus shengliensis*, two thermophilic methanogens belonging to different orders. Biochemical investigations combined with X-ray crystallography unveiled the first structures of archaeal glutamine synthetases and highlighted differences in their regulation. The enzyme from *M. thermolithotrophicus* is inactive in its resting state and employs 2-oxoglutarate as an on-switch. The 2-oxoglutarate acts as a sensor of cellular nitrogen deficiency, and its reported cellular concentration remarkably overlays with that required for the enzyme activation. Its binding to an allosteric pocket leads to the reconfiguration of the active site and promotes a catalytically competent state. The homolog from *M. shengliensis* does not harbor the 2-oxoglutarate

binding motif and, consequently, is 2-oxoglutarate insensitive. Instead, it is directly feedback-inhibited by glutamine, as shown for bacterial homologs. The glutamine inhibition depends on a key arginine residue from the Asp50'-loop. The arginine is substituted by a glycine in *M. thermolithotrophicus,* abolishing the inhibitory effect. While the effectors are surprisingly different, the molecular switch controlling the glutamine synthetase activity is fundamentally the same and depends on the correct positioning of the Asp50'-loop and a catalytic arginine. Residue conservation suggests that both regulation mechanisms are widespread and not mutually exclusive across archaea.

Main Text

Introduction

Nitrogen is an essential component of all living cells. Its most reduced state, ammonia (here representative for NH_3 and NH_4^+ in equilibrium), is one of the most common nitrogen sources assimilated by the microbial world [1-4]. Ammonia enters the central nitrogen metabolism via two systems: the glutamine synthetase – glutamate synthase (GS-GOGAT) couple and/or the glutamate dehydrogenase (GDH). While the GDH carries out the reversible reaction of reductive amination of 2-oxoglutarate (2OG) using NAD(P)H, the couple GS/GOGAT performs an ATP-dependent process [5-8]. Because of its higher affinity for ammonia, the GS-GOGAT couple is notably known to be more effective than the GDH in ammonia-limited environments [1]. The nitrogen assimilation by GS-GOGAT is operated in two steps (Fig. S1A). The GS initially produces glutamine via the condensation of ammonia on glutamate, a reaction coupled to ATP-hydrolysis that requires Mg^{2+} or Mn^{2+} [1]. Then, GOGAT performs the deamination of the synthesized glutamine, and the ammonia transfer to 2OG to ultimately form two molecules of glutamate, with the concomitant oxidation of an electron donor (e.g., NADH) [1, 7].

GS can be categorized into three types [1, 9]. Type I is present in prokaryotes, and some homologs were found in eukaryotes. It organizes as a homo-dodecamer of ~55 kDa-large subunits [10]. Type II is composed of a homo-decameric assembly of ~40 kDa-large subunits and is common in bacteria and eukaryotes [11, 12]. Type III is found in bacteria, archaea and eukaryotes and is composed of a homo-dodecamer of ~75 kDa-large subunits [1, 13]. The GS type I (GSI) is further subdivided into three classes: (i) the GSI-α found in archaea*, Actinobacteria, Desulfobacterota* and *Bacillota* (formerly *Firmicutes*), (ii) the GSI-β, present in many bacteria and a few archaea and (iii) the GSI-γ mostly found in bacteria [9]. Studies on the evolution of this ancient enzyme led to different scenarios, explaining the puzzling phylogeny of GS that should be the result of gene deletion and horizontal gene transfer, including that GSI-α and GSI-β were already separated in LUCA and a loss of GSI-α and/or GSI-β occurred in different lineages [9, 14-17]. The GSI-α in *Bacillota* was obtained through lateral gene transfer from an archaeal ancestor [9, 14]. The structural characterization of some GSI-α and GSI-β unveiled a similar organization [6,

65 [18, 19]. The dodecamer is organized as two stacked hexameric rings, and each polypeptide is organized
66 into two parts, a shorter N-terminal domain (β-Grasp fold) and the remaining C-terminal domain (Fig.
67 S1B-G). The N-terminal domain is mainly responsible for the ring association, while the C-terminal
68 domain composes the main part of the ring and is responsible for inter-ring interactions. GSI-β differs
69 structurally from GSI-α by a 25 amino acid-long extension involved in inter-ring stabilization (Fig. S1B-
70 D). The active sites are located at the interface of the N- and C- terminal domains of the adjacent subunit
71 in the hexameric ring. Each active site is structured as a "bifunnel" with ATP and glutamate binding on
72 opposite sides. The ATP binding site is usually referred to as the top of the funnel since the opening lies
73 towards the external ring surface, and the glutamate-binding site as the bottom of the funnel. Divalent
74 metal cations (Mn^{2+} or Mg^{2+}) are coordinated on the C-terminal domain and positioned at the center of
75 the bifunnel. Two loops in the active center are critical for catalysis: the Glu-flap, involved in shielding
76 the active site during catalysis and deprotonating the intermediate product, and the Asp-50' loop, which
77 binds and deprotonates ammonium. In this concerted orchestra, the glutamate is first phosphorylated on
78 its γ-carboxyl group by the ATP donor, and secondly, the ammonia is incorporated, releasing the
79 products glutamine, ADP and inorganic phosphate [6, 20]. GSI-γ are comprised of only the catalytic domain
80 lacking the N-terminus. A majority of biochemically characterized GSI-γ members lack GS biosynthetic
81 activity and instead function as γ-glutamyl-polyamine synthetases [9].

82 The GSI-α and GSI-β also differ regarding their regulation. The GSI-β have a complex multi-level
83 regulation by adenylylation, a feature not conserved in GSI-α [8]. The regulation of GSI-α in *Bacillus subtilis*
84 was shown to depend on feedback inhibition by the enzyme's product glutamine [18]. The glutamine-
85 inhibited enzyme also binds the regulator GlnR, triggering a switch from a dodecameric GS to an inactive
86 tetradecameric form [19]. The GS-GlnR complex finally acts as a transcription repressor for genes involved
87 in nitrogen assimilation, including the GS.

88 The archaeal GSI-α has never been structurally characterized, and only sporadic studies exist [21-25]. While
89 the activity of the GSI-α of *Haloferax mediterranei* is inhibited by glutamine [26], other radically different
90 regulations have been discovered. In the same archaeon, the molecule 2OG was shown to be a more
91 efficient regulator with a 12-fold activity stimulation, reaching 18-fold when P_{II}-family regulatory proteins
92 $GlnK_1$ or $GlnK_2$ were added in addition to 2OG [26]. 2OG and GlnK also enhance the activity of the enzyme
93 from the methanogen *Methanosarcina mazei*, as well as the 23 residue-large sP26 protein [27, 28], although
94 an inhibitory effect of GlnK (under different buffer conditions) has also been shown [28]. The regulatory
95 proteins GlnK and sP26, whose expression is triggered under nitrogen starvation, form a complex and
96 make a tight and specific interaction with the GS [27, 29]. The direct control by 2OG is particularly elegant
97 as the metabolite is a cellular sensor for nitrogen deprivation and has been shown to act on several key
98 nodes of nitrogen acquisition in methanogens (Fig. S1A) [8, 26-28, 30].

99 To shed light on the regulatory mechanisms of archaeal GSI-α, we performed biochemical and structural
100 characterizations on the enzymes from two thermophilic methanogens belonging to different phyla. Our
101 results illustrate the different regulation systems at the molecular level, including an unforeseen switch-
102 on controlled by 2OG.

103

104 **Results**

105 ***Mt*GS activity is strictly dependent on 2OG in contrast to *Ms*GS.**
106 *Methanothermococcus thermolithotrophicus* belonging to the *Methanococcales* order is a strictly
107 hydrogenotrophic methanogen which possesses a single gene coding for a GSI-α (the gene product is
108 referred to as *Mt*GS) while lacking any genes coding for GDH. The purification of *Mt*GS was previously
109 published [31]. We improved the purification protocol to obtain an anaerobically purified *Mt*GS allowing its
110 detailed enzymatic characterization. Denaturing and high-resolution clear native polyacrylamide gel
111 electrophoresis (hrCN PAGE) are coherent with a dodecameric organization of the ~50 kDa peptide (Fig.
112 1A and B). The activity of *Mt*GS, measured via a coupled enzyme assay (Fig. S2), could not be detected
113 despite the addition of Mg^{2+} or Mn^{2+}, an increase of the protein concentration (up to 0.75 mg/ml final
114 concentration) or an extension of the incubation time. As 2OG acts as an activity stimulator in archaeal
115 GS [26-28], 2OG was added to the assay and revealed to be essential for *Mt*GS activity under these
116 experimental conditions (Fig. 1C). No activity was detected below 60 µM 2OG, 50% activity was reached
117 at 170 µM and saturation occurred above ~ 0.6 mM. Noteworthy, exposure to oxygen severely reduced
118 the enzyme activity (1 h of exposure to O_2 decreased the activity to 6%). Specificity for 2OG was tested
119 by substitution with malate or succinate, which are structurally similar to 2OG. At 15 mM, both surrogates
120 could not activate *Mt*GS (Fig. S3). However, reactivation by 2 mM 2OG was still observed (Fig. S3).

121 To investigate whether the 2OG dependency is a feature unique to *M. thermolithotrophicus* or also
122 distributed among other methanogens, we purified the GS from the phylogenetically distant methanogen
123 *Methermicoccus shengliensis*. This archaeon belongs to the *Methanosarcinales* order and uses
124 methylotrophic methanogenesis for energy and carbon acquisition [32]. Its genome codes for a single GSI-
125 α (*Ms*GS) and no GDH. *Ms*GS was anaerobically purified from *M. shengliensis* leading to a major band
126 at ~48 kDa on SDS PAGE (expected molecular weight 49 kDa, Fig. 1A). A dodecameric assembly was
127 confirmed by hrCN PAGE (Fig. 1B). Unlike for *Mt*GS, 2OG was not required for enzyme activity, and its
128 addition (2 mM) did not affect its activity (Fig 1D). Kinetic parameters of both enzymes were determined
129 and are summarized in Fig. 1E. Notably, *Mt*GS reaches a 30-fold higher apparent V_{max} than *Ms*GS once
130 fully activated with 2OG. Both enzymes exhibit similar apparent K_m for glutamate and ammonium, and
131 binding of both substrates exhibits slight positive cooperativity in *Mt*GS, while in *Ms*GS, only glutamate
132 binding was cooperative.

133 Structural studies were undertaken to investigate the activation mechanism and the difference between
134 both GS systems.

135 **The first structures of archaeal GS**

136 The structure of *Mt*GS in its resting state was solved by using the diffraction data obtained by Engilberge
137 *et al.* [31]. The data was collected on a crystal obtained by co-crystallization with a crystallophore (TbXo4)
138 that corrected the twinning issue inherent to this crystalline form. The structure (*Mt*GS-apo-TbXo4) was
139 solved by molecular replacement using the closest structural homolog from *Bacillus subtilis* (*Bs*GS
140 resting state, PDB 4LNN [18]) and refined to a resolution of 1.65-Å (Table S1). The *Mt*GS structure presents
141 the typical homo-dodecameric architecture as seen in other structural homologs, with a monomeric unit
142 divided into Nter- (1-111) and Cter- (112-448) domains (Fig. 2A-C). A deviation of 0.847-Å (on 338 Cα
143 aligned from one monomer) exists between this structure and the apo structure of *Bs*GS (Fig. S4). The
144 TbXo4, contributing to the crystal packing, might have artificially provoked this deviation. Moreover,
145 Terbium atoms are located in the active site, coordinated by the residues involved in the Mg^{2+}/Mn^{2+}
146 recognition (Fig. S5). To exclude any artefactual effect of TbXo4, we performed crystallization in the
147 absence of the compound. The resulting crystal belonged to the same space group as the TbXo4-
148 containing form, but the crystallographic data exhibited a pseudomerohedral twinning with a fraction of
149 0.12. The protein structure (*Mt*GS-apo without TbXo4) was refined to 2.43-Å and is almost identical to
150 the TbXo4-containing form (RMSD of 0.206-Å, 410 atoms aligned Fig. S5, residues 67-69 could not be
151 modeled in the TbXo4-lacking structure), excluding the sequestration of an artefactual conformation
152 provoked by the compound. For this reason, the *Mt*GS TbXo4-containing structure obtained at the higher
153 resolution of 1.65 Å was used for further analyses.

154 The enzyme from *M. shengliensis* was crystallized in the absence of ligands, and two different crystalline
155 forms, referred to as *Ms*GS-apo 1 and *Ms*GS-apo 2, were analyzed (Table S2). *Mt*GS-apo-TbXo4 was
156 selected as a template for molecular replacement, and the structures were refined to 2.64-Å and 3.09-
157 Å, respectively (Fig. S6). As *Mt*GS, the homolog from *M. shengliensis* has a dodecameric organization
158 with the N-ter (1-106) and C-ter (107-442) domains forming the protomeric unit (Fig. 2D-F). *Ms*GS-apo
159 1 was used for the following structural analyses due to the better resolution.

160 A more detailed view of the monomeric structures highlights the key catalytic elements surrounding the
161 active site in both GS (Fig 2C and 2F, Fig S7 [18]): the Glu-flap (*Mt*GS 306-311, *Ms*GS 299-304), the Tyr-
162 loop (*Mt*GS 370-378, *Ms*GS 363-371), the Asn-loop (*Mt*GS 236-247, *Ms*GS 229-240), Tyr179-loop
163 (*Mt*GS 154-163, *Ms*GS 147-156) and Asp50'-loop (*Mt*GS 57-71, *Ms*GS 51-65). A structural comparison
164 of the monomeric unit reveals that *Ms*GS has a closer fit with *Bs*GS compared to *Mt*GS, which exhibits
165 larger deviations such as an extended helix α2 resulting from a seven-residue insertion (Fig 2G, Fig.
166 S4). While some deviations of local loops might be attributed to the increase in flexibility (Fig. S4A), both

167 archaeal GS deviate from *Bs*GS at the loop before helix α4 (highlighted by a star on Fig. 2G and 2H) at
168 the subunit interface.

169 The GS are well known to be dynamic, and structural rearrangement can occur upon ligand binding [6, 18,
170 19, 33, 34]. For instance, *Bs*GS undergoes dramatic intersubunit conformational movements between the
171 apo and transition state. Therefore, to investigate why *Mt*GS apo resting state would be catalytically
172 inactive, co-crystallization with 2OG was undergone.

173 **A 2OG allosteric site localized at the inter-subunit junction**

174 The complex with 2OG/Mg^{2+} was obtained with and without ATP (*Mt*GS-2OG/Mg^{2+}/ATP, *Mt*GS-
175 2OG/Mg^{2+}) in a new crystalline form and both were refined to 2.15-Å and 2.91-Å resolution, respectively
176 (Table S1). Both 2OG-containing structures show high tridimensional conservation (RMSD 0.279-Å, 441
177 atoms aligned, Fig. S8) while deviating from the apo resting state (2OG/Mg^{2+}/ATP: RMSD 0.694-Å, 374
178 atoms aligned, 2OG/Mg^{2+}: 0.631-Å, 371 atoms aligned Fig. S9 and S10). Since the 2OG/Mg^{2+}/ATP
179 complex is at a higher resolution, this structure was preferred for deeper structural analyses (Fig. S11).
180 The allosteric 2OG binding site, ~15 Å distant from the catalytic cleft, is localized at the interface between
181 the C-terminal domain and the adjacent N-terminal domain. 2OG is coordinated via ionic bonds by Arg20'
182 Arg88', Arg175, Arg174, and a hydrogen bond via Ser191 (primed numbers indicate the adjacent unit,
183 Fig 3A). Phe18' stabilizes the 2OG via stacking with its phenyl ring. In the apo structure, water molecules
184 replace 2OG in a pocket with a similar volume and slightly open to the solvent (Fig. S11C). In the 2OG-
185 insensitive *Ms*GS structure, the side chain of Glu89' substituting the *Mt*GS Val94' would collide with the
186 2OG position (Fig. 3B). Moreover, in *Ms*GS, most of the residues relevant for 2OG coordination (Arg88',
187 Arg174 and Ser191) are substituted. Interestingly, the structures of the bacterial homologs show a
188 tyrosine and glutamate side chain at the equivalent positions 18' and 34' in *Mt*GS that would collide with
189 the modeled 2OG (Fig. 3C). Additionally, the 2OG-coordinating Arg88' is absent (Fig. 3D). The Arg88'
190 might be the essential component for 2OG-based activation as it is positioned on the loop that undergoes
191 the most significant conformational change during 2OG binding. Therefore, its movement provoked by
192 2OG binding would trigger further movements throughout the polypeptide chain (see supplementary
193 movie "*Mt*GS_conformational_changes").

194 **2OG-dependent motions and active site remodeling in *Mt*GS.**

195 2OG presence at the interface of monomers reorients the bulky side chain of Phe18', and remodels salt
196 bridge networks in its surrounding. Consequently, the N-terminal domain is pushed away from the
197 neighboring C-terminal domain, and the α-helix 6 and β-hairpin 9-10 (residues 167-206, in the vicinity of
198 the active site) are shifted (Fig. 3E). The structural alignment on the C-terminal part from one monomer
199 exemplifies a 4.8° oscillation of the N-terminal domain, leading to a movement of 3-Å (Fig. S9). The
200 repositioning of the N-terminal domain would clash with the next monomer, and therefore a domino effect

201 occurs in the hexamer ring upon 2OG binding. This probably explains the cooperative behavior observed
202 in our kinetics data (Hill coefficient 3.4 for 2OG). Overall, the transition between *Mt*GS resting and 2OG-
203 bound states provokes a compaction of the outer ring and an opening in the center and middle parts
204 (Fig. 3F), leading to a more "opened" conformation. The observed switch and the 2OG-bound active
205 state are similar to the one observed between apo and complex of *B. subtilis*. The most dramatic change
206 occurs in the Asp50'-loop (critical for glutamate and ammonia coordination) interacting with the loop 318-
207 325, which reorganize to reach a position similar to that observed in the bacterial homologs co-
208 crystallized with the L-methionine-S-sulfoximine-phosphate (abbreviated as SOX, a mimic of a reaction
209 transition state) [18, 19] (Fig. S12).

210 In the *Mt*GS-2OG/Mg^{2+}/ATP structure, the ATP is located in a similar position compared to bacterial
211 homologs (Fig. 4A and B). The adenosine part is stacked in between Phe206 and Arg336 and
212 coordinated via hydrogen bonds with Lys333 main chain and Ser254 side chain. The ribose moiety is
213 coordinated by the Phe204 and Phe206 main chain and Glu189 side chain. Arg336, Arg326, and Arg321
214 bind the triphosphate backbone with ionic bonds. In the resting state, the ATP would clash with the
215 Phe206 and would not bind without displacing the β-sheet 10 (residues 199-205) and the following loop
216 (206-209) (Fig. 4C and S13).

217 We also obtained and refined to 2.70-Å a *Ms*GS structure in complex with ATP and Mg^{2+}. Only minor
218 differences exist compared to the apo state (Fig. S10). While the adenine part of the ATP is bound in a
219 similar way as described before, the ribose backbone is tilted by 90 °, pointing the triphosphate away
220 from the active site (Fig. 4D). This results in a non-catalytic state, similar to the one observed in the
221 *Bs*GS structure loaded with AMPPCP (PDB 4LNK [18]). Nevertheless, a sequence alignment of *Bs*GS,
222 *Mt*GS and *Ms*GS revealed a complete conservation of the residues involved in nucleotide binding (Fig
223 4E). Therefore, we concluded that the three GS share the same ATP coordination throughout the
224 catalytic steps.

225 2OG binding in *Mt*GS also impacts the glutamate binding site architecture. The residues involved in
226 glutamate binding (or the mimic SOX molecule) in *B. subtilis* are perfectly conserved in *Mt*GS, *Ms*GS
227 and other archaeal GS (Fig. 5). The conservation between the archaeal and bacterial domains reinforces
228 the importance of these residues to orchestrate the catalysis. While the predicted residues involved in
229 Mg^{2+}/glutamate binding share the same position between the *Mt*GS resting and 2OG-bound states
230 (Glu137, Glu139, Glu194, Gly246, His250, Glu309, Arg303, Arg340), the catalytic Arg321 (homologous
231 to Arg316 in *Bs*GS) shows a drastic difference of position (Fig. 5B and C). In the resting state, the loop
232 318-325 is constrained by a network of hydrogen bonds and salt bridges from the Asp50'-loop, the side
233 chain of the Arg321 itself being sequestered by Glu70'. Strikingly, in the 2OG-bound state of *Mt*GS, the
234 318-325 and Asp50´-loops adopt a position similar to all described active transition states in bacterial

235 homologs (Fig. S14). This suggests that the binding of 2OG triggers conformational changes leading to
236 a catalytically competent conformation and explaining the 2OG dependency for activity. Moreover, the
237 Glu309 from the Glu-flap is disengaged and in a more relaxed state in the 2OG-bound model, allowing
238 its assistance for catalysis. In *Ms*GS, the β-hairpin β3-β4 carrying the Asp50'-loop appears to adopt a
239 similar position with or without ATP. However, the Asp50' loop is flexible and cannot be modeled. The
240 flexibility might directly affect the stabilization of the adjacent loop 313-319 (318-325 in *Mt*GS) that
241 exhibits a different position than that found in *Mt*GS and bacterial GS with the catalytically important
242 Arg314 (Arg321 in *Mt*GS) retracted on the Tyr-loop (Fig. 5D).

243 **_Ms_GS is feedback inhibited by glutamine**

244 *Bs*GS is inhibited by the reaction product glutamine [18, 19], a feedback regulation under the control of the
245 Asp50' loop. When glutamine occupies the substrate-binding site, it interacts with Glu304 (*Bs*GS
246 numbering), required for catalysis. The Arg62' from the Asp50' loop stabilizes this interaction, resulting
247 in a lock of the Glu-flap, closing the active site. The Arg62' is not conserved in *Mt*GS (substituted by
248 Gly67') but is present in *Ms*GS (Arg61', Fig. 5E). To investigate this feedback inhibition mechanism, we
249 performed activity assays with glutamine addition for *Mt*GS and *Ms*GS. As suspected from the
250 sequences, the addition of glutamine (20 mM) inhibited the activity of *Ms*GS by 97%, while no difference
251 was observed for *Mt*GS (Fig. 5F). The substitution of the arginine appears to completely abolish
252 glutamine inhibition.

253 **Direct regulation by 2OG and glutamine within the domain *Archaea***

254 In addition to the previous studies of archaeal GSI-α [26-28], the characterization of the enzyme from *M.*
255 *thermolithotrophicus* and *M. shengliensis* highlighted that the mechanisms of post-translational
256 regulation are not conserved among archaea. Our study points out that the direct binding of 2OG and
257 glutamine depends on a restricted number of residues, suggesting that the potential regulation of GS
258 activity by both 2OG and glutamine might be predictable from the amino acid sequences. The
259 conservation of residues involved in regulator-binding was analyzed through 500 GS sequences
260 regrouping *Mt*GS and its closest archaeal homologs in the RefSeq database (Fig. 6 and Fig. S15). The
261 GSI-α sequences used to build the tree are, in most cases, forming monophyletic branches that separate
262 archaeal orders (sometimes different groups of GS within orders) and that apparently share a regulation
263 mechanism (see residue conservation in Fig. S15), except for some sequences including *Ms*GS which
264 do not clearly branch with any other GS (Fig. 6). Around half of the GS groups isolated from this tree (8
265 over the 17 archaeal GS groups) appear to harbor the residues necessary for binding either 2OG or
266 glutamine. In comparison, a third of the GS groups can be theoretically regulated by both (6 over the 17
267 archaeal GS groups, Fig. 6), of which the enzymes from *H. mediterranei* and *M. mazei* are
268 representatives. In certain archaeal orders (e.g. *Methanosarcinales*), the genome encodes two GS
269 isoforms that appear to have different sensitivity toward 2OG and glutamine. The enzymes found in

270 *Thermococcales*, *Thermoplasmatales* or *Methanomassiliicoccales* present substitutions that would
271 hinder their capacities to bind both 2OG and glutamine. Yet, binding at another position in the protein
272 structure cannot be excluded to allow potential interactions with regulatory partners. This analysis, yet
273 restricted to a small number of archaeal orders and therefore far from describing the overall archaeal
274 GS family, suggests variability in regulation strategies among these enzymes pointed out by the
275 enzymatic and structural analyses.

276

277 **Discussion**

278 In the absence of a GDH system, GS represents the main entry point for ammonia assimilation, feeding
279 the cellular metabolism with nitrogen. The ATP-dependent reaction must be under tight control to fit
280 cellular needs. The GS comes in different flavors in eukaryotes and prokaryotes, as does their regulation
281 [8, 18, 19, 26-28]. Many studies have been performed on bacterial and eukaryotic enzymes, but only a few are
282 focused on the GS from the domain *Archaea*. This work unveiled the first GSI-α structures of
283 representatives from archaea. Their homo-dodecameric organization is highly similar to their bacterial
284 counterparts. This is in accordance with published phylogenetic analyses in which archaeal and bacterial
285 GSI-α are part of a monophyletic group (Fig. 6 [9]). Our results expand the knowledge gathered on this
286 enzyme and aim to dissect the regulation mechanism at the molecular level.

287 2OG is a sensor for cellular nitrogen availability and has recently been recognized as a master regulator
288 of multiple biochemical pathways [30], especially other elements of nitrogen assimilation. For instance,
289 most of the P_{II}-family regulatory proteins, operating the signaling for nitrogen fluxes (including the
290 regulation of GS), are regulated by 2OG [30, 35-38]. Previous studies highlighted 2OG as a direct regulator
291 of GSI-α in *Archaea* [26-28], such as *Haloferax mediteraneii* and *Methanosarcina mazei*, in which 2OG
292 activates the enzyme by 12-fold [26] and 16-fold [28], respectively. Our study found that *Mt*GS activity is
293 strictly dependent on this metabolite under the described experimental conditions. *Mt*GS activation is
294 saturated at 0.6 mM 2OG and is inactive below 0.06 mM. This range remarkably fits the measured 2OG
295 cellular concentration in the methanogen *Methanococcus maripaludis*, which is around 0.8 mM under
296 N_2-fixing conditions (nitrogen-limited) and 0.08 mM after ammonium addition [39]. Relying on 2OG
297 concentration as a sensor of cellular nitrogen availability could benefit an energy-limited organism such
298 as *M. thermolithotrophicus*, and might represent a “primitive” regulation mechanism, which evolved
299 before more elaborate regulatory networks emerged. In addition, regulation at the transcriptional level
300 also occurs as *M. thermolithotrophicus* upregulates *glnA* expression when cells switch to N_2 fixation [40],
301 and it is known that small RNAs are involved in nitrogen metabolism regulation in prokaryotes, including
302 methanogens [41].

303 The structure of *Mt*GS reveals an allosteric pocket perfectly suited to accommodate 2OG via specific
304 salt bridges, hydrogen bonds, and Van der Waals contacts. Binding of 2OG leads to a succession of
305 conformational rearrangements resulting in a catalytically competent state. A conservation of all five
306 residues of the 2OG binding site can be found in other orders of archaea (Fig. 6 and S15). However,
307 while the motif would allow 2OG binding, the influence on the activity remains to be verified as several
308 other structural features are involved in the accurate positioning of catalytic residues. A 2OG binding
309 motif is far from being a feature shared by all archaea, as exemplified by the characterization of the 2OG-
310 insensitive *Ms*GS. Previous works on the enzymes from *H. mediterranei* and *M. mazei* pointed towards
311 a unified 2OG activation mechanism as main regulation in archaeal GS. However, spreading the analysis
312 to other species rather highlights a variety of different regulation mechanisms depending on archaeal
313 groups and enzyme isoforms (Fig. 6). The glutamine inhibition for *Ms*GS, previously described in *Bs*GS
314 [18], is dependent on one essential arginine (R62') responsible for the Glu-flap sequestration preventing
315 the release of the product from the active site, which is not happening in *Mt*GS due to a substitution
316 (Gly67').

317 Even if both, 2OG and glutamine, act through different mechanisms, e.g. competitive inhibition *versus*
318 allosteric activation, it is interesting that their modulation acts on the same key determinants. In the
319 resting state of *Mt*GS without 2OG or the glutamine feedback-inhibited *Bs*GS, the Glu-flap and the
320 catalytic arginine are locked in an unproductive conformation by the Asp50'-loop. In *Mt*GS the 2OG
321 binding provokes a major displacement of the N-terminal domain towards the adjacent C-terminal
322 domain and the overall conformational changes promote a restructuration of the Asp50'-loop, leading to
323 the repositioning Arg321 and of the Glu-flap, yielding an active conformation. Additionally, the 2OG-
324 dependent motions provide enough flexibility to allow the correct structuration of the Mg^{2+}/ATP binding
325 site, similar to *Ms*GS in its resting state.

326 Kinetic measurements revealed similar apparent K_M for ammonia and glutamate between both GS (Fig.
327 1E) and are comparable to previous studies [18, 42]. However, they differ more drastically in their apparent
328 V_{max}. While they are still in the range of previously reported specific activities of bacterial (up to 23
329 $\mu mol.min^{-1}.mg^{-1}$ for *B. subtilis* [43]) and archaeal glutamine synthetases (up to around 6 $\mu mol.min^{-1}.mg^{-1}$ in
330 both *H. mediterranei* and *M. mazei* [26, 28]), the apparent V_{max} of *Ms*GS is about 30 times lower compared
331 to *Mt*GS and might indicate the need for additional regulators. The sP26 protein described in *M. mazeii*
332 is not encoded in the *M. thermolithotrophicus* and *M. shengliensis* genomes. However, P_{II}-family proteins
333 might play a role in *Ms*GS activation that could mediate 2OG sensitivity, a hypothesis that future
334 exploratory studies will hopefully confirm.

335 The works gathered on GS illustrate that despite the universal requirement of assimilating ammonia to
336 fuel nitrogen metabolism and a conserved reaction mechanism, microbes rely on different ways to

337 modulate this ATP-dependent activity. Such regulatory networks might have been elaborated over time
338 to adapt to a particular physiology, environment, or catabolism. Opening scientific investigation to a
339 broader group of organisms will hopefully contribute to extending our knowledge of this crucial and
340 ancient enzyme.

341

342 **Materials and Methods**

343 **Cultivation**

344 *M. thermolithotrophicus* (DSM 2095) cells were obtained from the Leibniz Institute DSMZ - German
345 Collection of Microorganisms and Cell Cultures (Braunschweig, Germany). Cells were grown in a
346 minimal medium as described in Jespersen et al. [44]. Anaerobic cultivation of *M. thermolithotrophicus* was
347 performed in a fermenter with NH_4Cl and SO_4^{2-} as described in Jespersen et al. 2022 [45] with slight
348 modifications, or in a fermenter with N_2 and SO_4^{2-} as described in Maslać et al 2022 [40]. We did not
349 observe dramatic changes in the expression of GS nor its properties between both cultivation processes.
350 *M. shengliensis* ZC-1 (DSM 18856) was also obtained from the DSMZ (Braunschweig, Germany) and
351 was grown anaerobically on methanol as previously described in Kurth and Müller et al. 2021 [46].

352 **Purification**

353 Cell lysis and extracts were prepared in an anaerobic chamber at room temperature filled with an N_2/CO_2
354 atmosphere (90:10 %). Enzyme purification was carried out under anaerobic conditions in a Coy tent
355 filled with an N_2/H_2 atmosphere (97:3 %), at 20 °C and under yellow light. For each step, chromatography
356 columns were washed with at least three column volumes (CV) of the corresponding loading buffer, and
357 samples were filtrated on 0.2 µm filters (Sartorius, Germany) prior to loading. During purification, the
358 enzyme was followed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and
359 absorbance monitoring at 280 nm. Purifications were performed with adapting protocols, and the optimal
360 one is described below for *Mt*GS and *Ms*GS.

361 *Mt*GS purification

362 25 g (wet weight) cells were thawed and diluted in 200 ml lysis buffer (50 mM Tricine/NaOH pH 8, 2 mM
363 dithiothreitol (DTT)), sonicated (~5 x 10 sec at 60%, probe KE76 Bandelin SONOPULS, Germany) and
364 centrifuged at 45,000 x *g* for 45 min at 18 °C. The supernatant was passed twice on a 4x5 ml HiTrap™
365 DEAE Sepharose FF column (Cytiva, Sweden). Elution was performed with a NaCl gradient ranging
366 from 150 to 500 mM in the same buffer. The gradient was applied for 90 min at 2 ml/min, and *Mt*GS was
367 eluted at 170 to 260 mM NaCl. The pool was diluted 1:4 in 20 mM $NaPO_4^{2-}$ pH 7.6 and 2 mM DTT. The
368 sample was passed twice on 2 x Mini CHT ™ type I column (Bio-Rad, United States). Elution was
369 performed with a gradient of 0 to 200 mM $NaPO_4^{2-}$ in 10 min, followed by a second gradient from 200 to

370 500 mM $NaPO_4^{2-}$ in 10 min. The gradient was run at 1.5 ml/min. Under these conditions, *Mt*GS was
371 eluted at 20 to 200 mM $NaPO_4^{2-}$. The resulting pool was concentrated to 900 µl with a 30-kDa cutoff
372 concentrator (Sartorius, Germany). The sample was injected thrice on a Superose™ 6 Increase 10/300
373 GL column (Cytiva, Sweden). The protein was eluted at 0.4 ml/min in 25 mM Tris/HCl pH 7.6, 10%
374 glycerol, 150 mM NaCl, and 2 mM DTT. The resulting pool was diluted 1:1 in 25 mM Tris/HCl pH 7.6, 2
375 M $(NH_4)_2SO_4$, 2 mM DTT, and injected on a Source™15 Phe 4.6/100 PE column (Cytiva, Sweden).
376 Elution was performed with a gradient ranging from 1 to 0 M $(NH_4)_2SO_4$ for 60 min at 0.5 ml/min. *Mt*GS
377 was eluted between 0.86 to 0.74 M $(NH_4)_2SO_4$. The final pool was diluted 1:100 in 50 mM Tricine/NaOH
378 pH 8, 2 mM DTT, and loaded on a MonoQ™ 5/50 GL column (Cytiva, Sweden). *Mt*GS was eluted with
379 a NaCl gradient ranging from 0 to 200 mM NaCl in 10 min followed by 200 to 500 mM NaCl in 60 min at
380 0.5 ml/min (eluted at 375-395 mM NaCl). The final pool was washed 1:1000 in 25 mM Tris/HCl pH 7.6,
381 10% glycerol, 150 mM NaCl, 2 mM DTT and concentrated to 200 µl with a 30-kDa cutoff concentrator
382 (Sartorius, Germany). The protein concentration was determined via Bradford assay, and the sample
383 was flash-frozen and stored anaerobically at -80 °C.

384 *Ms*GS purification

385 The pellet (4 g) was suspended in 50 mM Tris/HCl pH 8 and 2 mM DTT (lysis buffer). Cell lysis and
386 preparation of extracts were performed similarly, except that the pellet after centrifugation was
387 resuspended in the lysis buffer, sonicated and centrifuged a second time to extract additional protein.
388 Soluble fractions were pooled and diluted 1:1.4 with the lysis buffer (a total 15-fold dilution was reached
389 to decrease the salt concentration). The filter sample was loaded on a 2 x 5 mL HiTrap™ DEAE
390 Sepharose FF (Cytiva, Sweden) equilibrated with the same buffer. The protein was eluted with a 0 to
391 400 mM NaCl linear gradient for 150 min at a 2 ml/min flow rate. *Ms*GS eluted between 240 mM and
392 310 mM NaCl. The pooled sample was diluted with three volumes of lysis buffer and was loaded on a 5
393 ml HiTrap Q HP™ column (Cytiva, Sweden). The protein was eluted with a 150 to 500 mM NaCl linear
394 gradient using a 70 min at a 1 ml/min flow rate. *Ms*GS was eluted between 0.38 and 0.41 M NaCl under
395 these conditions. The pooled sample was diluted with 1 volume of 25 mM Tris/HCl pH 7.6, 2 M $(NH_4)_2SO_4$
396 and 2 mM DTT before loading on a 5 ml HiTrap™ Phenyl Sepharose HP column (Cytiva, Sweden)
397 equilibrated with the same buffer. *Ms*GS was eluted with a 1.1 to 0 M $(NH_4)_2SO_4$ linear gradient for 70
398 min at a 1 ml/min flow rate. The protein was eluted between 0.91 M and 0.81 M $(NH_4)_2SO_4$. Pooled
399 fractions were concentrated on a 10-kDa cutoff centrifugal concentrator (Sartorius, Germany), and the
400 buffer was exchanged for 25 mM Tris/HCl pH 7.6, 10% glycerol, and 2 mM DTT.

401 5-12% gradient hrCN PAGE was adapted as described in Lemaire et al. [47] (originally described in
402 Lemaire et al. [48]) and run at 40 mA for 1 h.

403 **Activity assays**

404 GS activities were measured using the pyruvate kinase/ lactate dehydrogenase (PK/LDH) coupled
405 enzymes from rabbit muscle ordered from Sigma-Aldrich (containing 600-1,000 units/ml pyruvate kinase
406 900-1,400 units/ml lactate dehydrogenase). The activity was measured by following NADH oxidation
407 resulting in the change of absorbance at 340 nm (Fig. S2).

408 Absorbance was measured aerobically in a 96-well plate with a SPEKTROstarNano (BMG Labtech,
409 Germany) at 45 °C in 100 µl final reaction volume. GS was added immediately after opening the
410 anaerobic storage flask to minimize the negative effects of oxygen and low temperature. Activities were
411 measured in triplicate.

412 All reagents were prepared in the reaction buffer (200 mM KH_2PO_4 pH 7, 10 mM KCl) apart from $MgCl_2$
413 x $6H_2O$ which was dissolved in water. The standard reaction mix contained as a final concentration: 1
414 mM NADH (freshly prepared), 2 mM PEP (freshly prepared), 3 mM ATP (freshly prepared), 50 or 80 mM
415 sodium glutamate (for *Mt*GS and *Ms*GS, respectively), 20 mM NH_4Cl, 25 mM $MgCl_2$ x $6H_2O$, 2 mM
416 sodium 2OG (freshly prepared), 0.02 mg/ml GS and 5 µl PK/LDH from a 1:10 dilution of the stock. Kinetic
417 parameters were determined by varying the glutamate concentrations from 0 to 100 mM for *Mt*GS and
418 0 to 200 mM for *Ms*GS, and NH_4Cl from 0 to 5 mM for *Mt*GS and 0 to 20 mM for *Ms*GS. The determination
419 of the $K_{0.5}$ and hill coefficient for *Mt*GS activation was performed by varying the 2OG concentration from
420 0 to 5 mM. Glutamine feedback inhibition was measured at 20 mM glutamine. However, it has to be
421 mentioned that in this case, both *Mt*GS and *Ms*GS have been frozen at -80 °C for an extensive period
422 (2.5 and 1 year, respectively) which affected the specific activity. Unspecific interactions with malate and
423 succinate for *Mt*GS were measured at 15 mM of the compounds by preparing the reaction mix without
424 2OG, measuring for 20 min and then adding 2OG (2 mM final). Here, the activity was also recorded from
425 the enzyme frozen for one year at -80 °C. O_2 sensitivity was measured in anaerobic cuvettes with 0.01
426 mg/ml *Mt*GS, starting the reaction with glutamate addition. All solutions were prepared anaerobically,
427 and then half of the master mix, GS and glutamate were separately incubated for one hour at ambient
428 air. The anaerobic or aerobic compounds were combined, and activity was measured immediately. It
429 has been previously observed that certain buffer conditions contribute to a faster activity loss, including
430 compounds like β-mercaptoethanol or dithiothreitol [42]. Therefore, it should be considered when
431 comparing specific activities from different studies, as it might also affect the O_2-effect on enzymes.

432 Rates were measured by exploiting the linear regression for each dataset through the points with the
433 steepest, most linear slope. The control slope without enzyme was subtracted from the slope obtained
434 with the enzyme in the same timeframe. Activities are presented in µmol of NADH oxidized per minute
435 per mg of added GS, using a molar extinction coefficient of ε_{340nm}= 6,220 $M^{-1}.cm^{-1}$ for NADH. The
436 apparent K_M and V_{max} were calculated with the Km Vmax Tool Kit (ic50tk/kmvmax.html).

437 **Crystallization**

438 All proteins were crystallized fresh without any freezing step, and all were obtained through the sitting
439 drop method on a 96-Well MRC 2-Drop Crystallization Plates in polystyrene (SWISSCI, United Kingdom)
440 at 20 °C under anaerobic conditions (N_2:H_2, gas ratio of 97:3).

441 *Mt*GS apo without TbXo4 was crystallized at a concentration of 15 mg/ml. The crystallization reservoir
442 contained 90 µL of mother liquor (20 % w/v polyethylene glycol 3,350 and 100 mM potassium sodium
443 tartrate), the crystallization drop contained a mixture of 0.55 µL protein and 0.55 µL precipitant.
444 Hexagonal plates appeared after few weeks and were soaked in the mother liquor supplemented with
445 25% v/v ethylene glycol prior to freezing in liquid nitrogen.

446 *Mt*GS with 2OG/Mg^{2+}/ATP was crystallized at 3.7 mg/ml with a final concentration of 2 mM 2OG, 2 mM
447 ATP and 2 mM MgCl. The crystallization reservoir contained 90 µl of mother liquor (20 % w/v
448 polyethylene glycol 3,350 and 200 mM sodium fluoride). The crystallization drop contained 0.6 µl GS
449 with ligands and 0.6 µl precipitant. Short thick hexagonal rods appeared after few weeks and were
450 soaked in the mother liquor supplemented with 20 % v/v glycerol prior to freezing in liquid nitrogen.

451 *Mt*GS with 2OG/Mg^{2+} was crystallized at 3.4 mg/ml with a final concentration of 2 mM 2OG and 2 mM
452 MgCl. The crystallization reservoir contained 90 µl of mother liquor (25 % w/v Polyethylene glycol 1,500
453 and 100 mM SPG (succinic acid, sodium dihydrogen phosphate, and glycine) buffer pH 5.0), and the
454 crystallization drop contained 0.6 µl GS with ligands and 0.6 µl precipitant. Hexagonal plates obtained
455 after few weeks were soaked in the mother liquor supplemented with 15 % v/v glycerol prior to freezing
456 in liquid nitrogen.

457 All *Ms*GS crystals were obtained by mixing of 0.55 µL protein at 9 mg/ml with 0.55 µL precipitant.
458 For *Ms*GS-apo 1 the crystallization reservoir contained 90 µl of 200 mM ammonium formate and 20 %
459 (w/v) polyethylene glycol 3,350. Rectangular rods appeared within weeks, and crystals were soaked in
460 the mother liquor supplemented with 25% v/v ethylene glycol prior to freezing in liquid nitrogen. *Ms*GS-
461 Mg^{2+}/ATP was obtained from the same condition. Here, crystals were soaked for 4 min in the
462 crystallization solution supplemented with 10 mM ATP/Mg^{2+}/sodium glutamate and then back soaked in
463 the crystallization solution supplemented with 25% glycerol prior to freezing in liquid nitrogen. It is worth
464 noting that no additional electron density could be attributed despite soaking the *Ms*GS crystal with 10
465 mM glutamate. Therefore, the structure was named *Ms*GS-Mg^{2+}/ATP bound complex. The absence of
466 binding might come from a packing artefact, a low occupancy of magnesium in the active site, or the
467 aberrant position of the ATP phosphate backbone.

468 For *Ms*GS-apo 2 the crystallization reservoir contained 90 µl of 1.6 M sodium citrate tribasic dihydrate.
469 Flat squares appeared within weeks, and crystals were soaked in the mother liquor supplemented with
470 30% v/v glycerol prior to freezing in liquid nitrogen.

471 **Data collection, processing and structure refinement**

472 All datasets were collected at 100 K at the Swiss Light Source (SLS), beamline PXIII (X06DA). Data
473 processing was performed with *autoPROC* [49] combined with STARANISO [50] except for *Mt*GS-apo co-
474 crystallized with the TbXo4 (obtained from Engilberge et al. 2019 [31]), which was processed by XDS [51]
475 and scaled with SCALA from the CCP4 package. The structure of *Mt*GS-apo-TbXo4 was solved by
476 molecular replacement with MOLREP (CCP4) [52] by using the closest structural homolog from *Bacillus*
477 *subtilis* (*Bs*GS resting state, PDB 4LNN). Other *Mt*GS structures were solved by molecular replacement
478 with PHASER from the PHENIX package with the *Mt*GS-apo-TbXo4 model. *Mt*GS-2OG/Mg^{2+}/ATP was
479 used as template to solve *Mt*GS-2OG/Mg^{2+} and *Ms*GS-apo 1 structures with PHASER. The two other
480 *Ms*GS structures were solved with PHASER by using *Ms*GS-apo 1 as template.

481 Refinement was performed with *PHENIX* [53] and *BUSTER* [54] in combination with fast automatic visual
482 model building in *COOT* [55]. All models were systematically validated by using Molprobity [56]. *Mt*GS-apo
483 cocrystallized with TbXo4 was refined with considering all atoms except water anisotropic. Additionally,
484 riding hydrogens were added during the refinement. *Mt*GS-apo without TbXo4 was refined by applying
485 translation libration screw model (TLS) and adding riding hydrogens. Hydrogens were omitted in these
486 final deposited models. In addition, this dataset was refined by applying the following twin operator:
487 1/2h+1/2k,3/2h-1/2k,-l. *Mt*GS-2OG/Mg^{2+}/ATP was refined in BUSTER by applying non crystallography
488 symmetry (NCS) and TLS, while *Mt*GS-2OG/Mg^{2+} was refined in PHENIX with TLS and without NCS
489 and with riding hydrogens.

490 All *Ms*GS models were refined by applying TLS, without NCS and without generating hydrogens.

491 The data collection and refinement statistics for the deposited models are listed in Tables S1 and S2. All
492 models were deposited with the Worldwide Protein Data Bank (wwPDB) under the accession numbers
493 8OOL, 8OON, 8OOO, 8OOQ, 8OOW, 8OOX and 8OOZ,

494 **Structural analyses.**

495 All structure visualization was performed with PyMol (2.2.0 Schrödinger, LLC, New York, USA). RMSD
496 putty graphs were generated by generating RMSD values with SUPERPOSE from the CCP4 package
497 [52], substituting the B-values in the according PDB file with the new RMSD values and displaying it in the
498 Putty preset in PyMol. Coloring was set to “spectrum b, yellow_orange_magenta, minimum=0,
499 maximum=5”. Putty scaling was set to absolute linear scaling from the B factor column. The movie
500 visualizing the conformational changes upon 2OG binding was generated with UCSF Chimera (Version
501 1.16, University of California, USA). A morph between *Mt*GS-apo-TbXo4 to *Mt*GS-2OG/$Mg^{2+}$to *Mt*GS-

502 2OG/Mg^{2+}/ATP is shown. The enzyme is displayed as cartoon with the ligands as spheres. 2OG carbons
503 are colored yellow and ATP carbons light blue, while the other atoms are colored according to element
504 (O in red, N in blue, P in orange). In the second half a zoom-in of the active site is shown with the Asp-
505 50' loop in a darker blue and Arg321 is displayed as sticks. The two chains composing the active site
506 are colored in different shades of cyan.

507 Sequence alignments were generated with Clustal Omega [57] and superposed to the secondary structure
508 with the Espript 3.0 server (https://espript.ibcp.fr) [58].

509 **Phylogeny and residues conservation analysis.**
510 The *Mt*GS sequence was used as query for a BLAST P [59] search in the RefSeq database, restricted to
511 the *Archaea* domain. This database was used to limit the redundancy of sequences from similar species.
512 The research limit was set to 500 sequences. The sequences of the GS from *B. subtilis*, *S. aureus*, *P.*
513 *polymyxa* and *L. monocytogenes* were added, leading to a total of 504 sequences, all listed in Table S3.
514 The phylogenic tree was constructed by the MEGA X program [60]. The alignment was done by MUSCLE
515 with default parameters and the tree was constructed using the Maximum Likelihood method and JTT
516 matrix-based model [61]. The tree with the highest log likelihood (-191831.27) is shown. The node scores
517 were calculated with 200 replications. There were a total of 656 positions in the final dataset. The
518 branches were manually colored according to the taxonomy from the NCBI database [62]. The tree was
519 visualized and imaged using iTol v6 [63].
520 The sequences were manually separated in different monophyletic groups. The different sequences
521 forming the groups were aligned by Clustal Omega [57] and the residue conservation images were
522 constructed using Weblogo 3 (version 3.7.12) [64] based on the residues position in *Mt*GS and multiple
523 sequences alignments.
524

525 **Acknowledgments**
526 We would like to thank the Max Planck Institute for Marine Microbiology and the Max-Planck-Society for
527 their continuous support. We are also grateful for the beam time allocation at SLS and the support of the
528 PXIII beamline staff. We would also like to thank Christina Probian and Ramona Appel from the Microbial
529 Metabolism laboratory for their continuous support, as well as Marion Jespersen and Nevena Maslać for
530 providing *M. thermolithotrophicus* cells. We acknowledge Sylvain Engilberge, Éric Girard, Olivier Maury
531 and François Riobé for their contribution to the *Ms*GS-TbXo4 structure. We are grateful to Stian Torset
532 and Tomás Alarcón Schumacher for their help with data formatting for the RMSD graph generation. C.W.
533 and J.M.K. were supported by the SIAM gravitation program (grant #024002002) granted by the
534 Netherlands Organisation for Scientific Research and the Ministry of Education, Culture and Science.

535 J.M.K was furthermore supported by the Deutsche Forschungs Gesellschafts (DFG) Grant KU 3768/1-
536 1.
537
538 **Author Contributions:** *Methanothermococcus thermolithotrophicus* cultivation was mainly performed
539 by M-C.M. *M. shengliensis* was cultivated by J.M.K. *Mt*GS and *Ms*GS were purified and crystallized by
540 M-C.M and O.N.L., respectively. Biochemical characterization and activity assays were performed by M-
541 C.M. X-ray data collection was performed by M-C.M., O.N.L and T.W. Data processing, model building,
542 structure refinement, validation and deposition were performed by M-C.M and T.W. Structures were
543 analyzed by M-C.M. and T.W. O.N.L. performed phylogenetic analyses. C.U.W. and T.W. acquired
544 funding to realize the project. The paper was written by M-C.M, O.N.L. and T.W. with contributions and
545 final approval of all co-authors.

546 **Competing Interest Statement:** Authors declare no competing interests.

547 **References**

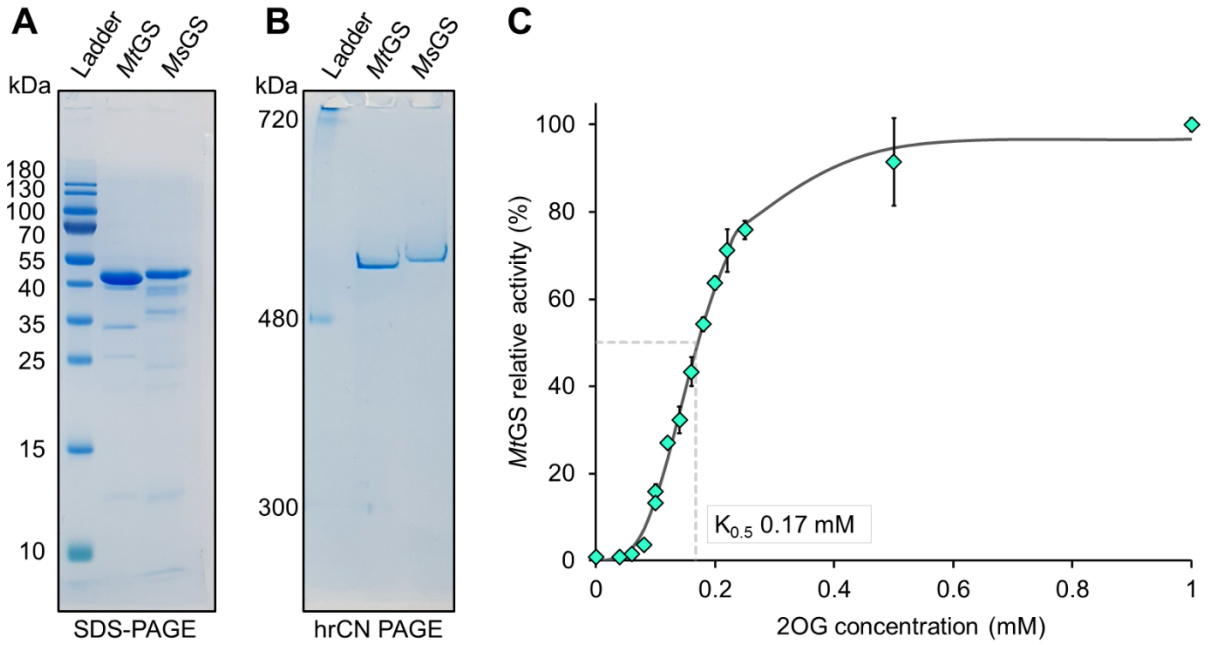
548 1. Cabello, P.; Roldán, M. D.; Moreno-Vivián, C., Nitrate reduction and the nitrogen cycle in
549 archaea. *Microbiology* **2004,** *150* (Pt 11), 3527-3546.
550 2. Kuypers, M. M. M.; Marchant, H. K.; Kartal, B., The microbial nitrogen-cycling network. *Nat Rev*
551 *Microbiol* **2018,** *16* (5), 263-276.
552 3. Zhang, X.; Ward, B. B.; Sigman, D. M., Global Nitrogen Cycle: Critical Enzymes, Organisms,
553 and Processes for Nitrogen Budgets and Dynamics. *Chem Rev* **2020,** *120* (12), 5308-5351.
554 4. Martinez-Espinosa, R. M., Microorganisms and Their Metabolic Capabilities in the Context of
555 the Biogeochemical Nitrogen Cycle at Extreme Environments. *International Journal of Molecular*
556 *Sciences* **2020,** *21* (12).
557 5. Arcondéguy, T.; Jack, R.; Merrick, M., P_{II} Signal Transduction Proteins, Pivotal Players in
558 Microbial Nitrogen Control. *Microbiol Mol Biol Rev* **2001,** *65* (1), 80-105.
559 6. Eisenberg, D.; Gill, H. S.; Pfluegl, G. M.; Rotstein, S. H., Structure-function relationships of
560 glutamine synthetases. *Biochim Biophys Acta* **2000,** *1477* (1-2), 122-45.
561 7. van den Heuvel, R. H. H.; Curti, B.; Vanoni, M. A.; Mattevi, A., Glutamate synthase: a
562 fascinating pathway from L-glutamine to L-glutamate. *Cellular and Molecular Life Sciences* **2004,** *61* (6),
563 669-681.
564 8. Leigh, J. A.; Dodsworth, J. A., Nitrogen Regulation in Bacteria and Archaea. *Annu Rev Microbiol*
565 **2007,** *61*, 349-77.
566 9. de Carvalho Fernandes, G.; Turchetto-Zolet, A. C.; Pereira Passaglia, L. M., Glutamine
567 synthetase evolutionary history revisited: Tracing back beyond the Last Universal Common Ancestor.
568 *Evolution* **2022,** *76* (3), 605-622.
569 10. Almassy, R. J.; Janson, C. A.; Hamlin, R.; Xuong, N. H.; Eisenberg, D., Novel subunit-subunit
570 interactions in the structure of glutamine-synthetase. *Nature* **1986,** *323* (6086), 304-309.
571 11. Unno, H.; Uchida, T.; Sugawara, H.; Kurisu, G.; Sugiyama, T.; Yamaya, T.; Sakakibara, H.;
572 Hase, T.; Kusunoki, M., Atomic Structure of Plant Glutamine Synthetase: A KEY ENZYME FOR PLANT
573 PRODUCTIVITY. *J Biol Chem* **2006,** *281* (39), 29287-96.
574 12. He, Y. X.; Gui, L.; Liu, Y. Z.; Du, Y.; Zhou, Y. Y.; Li, P.; Zhou, C. Z., Crystal structure of
575 *Saccharomyces cerevisiae* glutamine synthetase Gln1 suggests a nanotube-like supramolecular
576 assembly. *Proteins* **2009,** *76* (1), 249-254.
577 13. van Rooyen, J. M.; Abratt, V. R.; Belrhali, H.; Sewell, T., Crystal Structure of Type III Glutamine
578 Synthetase: Surprising Reversal of the Inter-Ring Interface. *Structure* **2011,** *19* (4), 471-483.

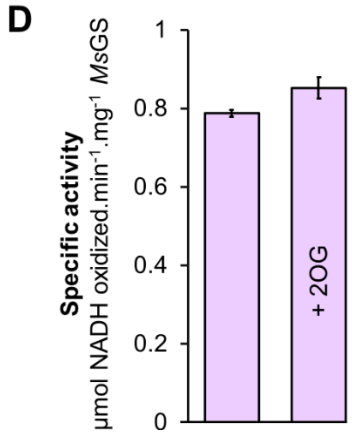
14. Pesole, G.; Gissi, C.; Lanave, C.; Saccone, C., Glutamine Synthetase Gene Evolution in Bacteria. *Mol Biol Evol* **1995,** *12* (2), 189-97.
15. Brown, J. R.; Masuchi, Y.; Robb, F. T.; Doolittle, W. F., Evolutionary Relationships of Bacterial and Archaeal Glutamine Synthetase Genes. *J Mol Evol* **1994,** *38* (6), 566-76.
16. Kumada, Y.; Benson, D. R.; Hillemann, D.; Hosted, T. J.; Rochefort, D. A.; Thompson, C. J.; Wohlleben, W.; Tateno, Y., Evolution of the glutamine-synthetase gene, one of the oldest existing and functioning genes. *P Natl Acad Sci USA* **1993,** *90* (7), 3009-3013.
17. Pesole, G.; Bozzetti, M. P.; Lanave, C.; Preparata, G.; Saccone, C., Glutamine synthetase gene evolution: A good molecular clock. *P Natl Acad Sci USA* **1991,** *88* (2), 522-526.
18. Murray, D. S.; Chinnam, N.; Tonthat, N. K.; Whitfill, T.; Wray, L. V., Jr.; Fisher, S. H.; Schumacher, M. A., Structures of the *Bacillus subtilis* Glutamine Synthetase Dodecamer Reveal Large Intersubunit Catalytic Conformational Changes Linked to a Unique Feedback Inhibition Mechanism. *J Biol Chem* **2013,** *288* (50), 35801-11.
19. Travis, B. A.; Peck, J. V.; Salinas, R.; Dopkins, B.; Lent, N.; Nguyen, V. D.; Borgnia, M. J.; Brennan, R. G.; Schumacher, M. A., Molecular dissection of the glutamine synthetase-GlnR nitrogen regulatory circuitry in Gram-positive bacteria. *Nat Commun* **2022,** *13* (1), 3793.
20. Liaw, S. H.; Kuo, I. C.; Eisenberg, D., Discovery of the ammonium substrate site on glutamine synthetase, a third cation binding site. *Protein Science* **1995,** *4* (11), 2358-2365.
21. Martinez-Espinosa, R. M.; Esclapez, J.; Bautista, V.; Bonete, M. J., An octameric prokaryotic glutamine synthetase from the haloarchaeon *Haloferax mediterranei*. *FEMS Microbiol Lett* **2006,** *264* (1), 110-6.
22. Adul Rahman, R. N.; Jongsareejit, B.; Fujiwara, S.; Imanaka, T., Characterization of Recombinant Glutamine Synthetase from the Hyperthermophilic Archaeon *Pyrococcus* sp. strain KOD1. *Appl Environ Microbiol* **1997,** *63* (6), 2472-6.
23. Neelon, K.; Schreier, H. J.; Meekins, H.; Robinson, P. M.; Roberts, M. F., Compatible solute effects on thermostability of glutamine synthetase and aspartate transcarbamoylase from *Methanococcus jannaschii*. *BBA-Proteins Proteom* **2005,** *1753* (2), 164-173.
24. Robinson, P.; Neelon, K.; Schreier, H. J.; Roberts, M. F., β-Glutamate as a Substrate for Glutamine Synthetase. *Appl Environ Microbiol* **2001,** *67* (10), 4458-63.
25. Yin, Z. M.; Purschke, W. G.; Schafer, G.; Schmidt, C. L., The Glutamine Synthetase from the Hyperthermoacidophilic Crenarcheon *Sulfolobus acidocaldarius*: Isolation, Characterization and Sequencing of the Gene. *Biol Chem* **1998,** *379* (11), 1349-1354.
26. Pedro-Roig, L.; Camacho, M.; Bonete, M. J., Regulation of ammonium assimilation in *Haloferax mediterranei*: Interaction between glutamine synthetase and two GlnK proteins. *Biochim Biophys Acta* **2013,** *1834* (1), 16-23.
27. Gutt, M.; Jordan, B.; Weidenbach, K.; Gudzuhn, M.; Kiessling, C.; Cassidy, L.; Helbig, A.; Tholey, A.; Pyper, D. J.; Kubatova, N.; Schwalbe, H.; Schmitz, R. A., High complexity of Glutamine synthetase regulation in *Methanosarcina mazei*: Small protein 26 interacts and enhances glutamine synthetase activity. *FEBS J* **2021,** *288* (18), 5350-5373.
28. Ehlers, C.; Weidenbach, K.; Veit, K.; Forchhammer, K.; Schmitz, R. A., Unique mechanistic features of post-translational regulation of glutamine synthetase activity in *Methanosarcina mazei* strain Gö1 in response to nitrogen availability. *Mol Microbiol* **2005,** *55* (6), 1841-54.
29. Ehlers, C.; Grabbe, R.; Veit, K.; Schmitz, R. A., Characterization of $GlnK_1$ from *Methanosarcina mazei* strain Go1: Complementation of an *Escherichia coli* gInK mutant strain by $GlnK_1$ (vol 184, pg 1040, 2002). *Journal of Bacteriology* **2002,** *184* (8), 2331-2331.
30. Huergo, L. F.; Dixon, R., The Emergence of 2-Oxoglutarate as a Master Regulator Metabolite. *Microbiol Mol Biol Rev* **2015,** *79* (4), 419-35.
31. Engilberge, S.; Wagner, T.; Santoni, G.; Breyton, C.; Shima, S.; Franzetti, B.; Riobé, F.; Maury, O.; Girard, E., Protein crystal structure determination with the crystallophore, a nucleating and phasing agent. *J Appl Crystallogr* **2019,** *52* (Pt 4), 722-731.
32. Cheng, L.; Qiu, T. L.; Yin, X. B.; Wu, X. L.; Hu, G. Q.; Deng, Y.; Zhang, H., *Methermicoccus shengliensis* gen. nov., sp. nov., a thermophilic, methylotrophic methanogen isolated from oil-production water, and proposal of *Methermicoccaceae* fam. nov. *Int J Syst Evol Microbiol* **2007,** *57* (Pt 12), 2964-2969.

633 33. Liaw, S. H.; Eisenberg, D., Structural Model for the Reaction Mechanism of Glutamine
634 Synthetase, Based on Five Crystal Structures of Enzyme-Substrate Complexes. *Biochemistry* **1994,** *33*
635 (3), 675-81.
636 34. Krajewski, W. W.; Collins, R.; Holmberg-Schiavone, L.; Jones, T. A.; Karlberg, T.; Mowbray,
637 S. L., Crystal Structures of Mammalian Glutamine Synthetases Illustrate Substrate-Induced
638 Conformational Changes and Provide Opportunities for Drug and Herbicide Design. *Journal of Molecular*
639 *Biology* **2008,** *375* (1), 217-228.
640 35. Arcondéguy, T.; Jack, R.; Merrick, M., P_{II} Signal Transduction Proteins, Pivotal Players in
641 Microbial Nitrogen Control. *Microbiol Mol Biol Rev* **2001,** *65* (1), 80-105.
642 36. Forchhammer, K.; Selim, K. A.; Huergo, L. F., New views on PII signaling: from nitrogen sensing
643 to global metabolic control. *Trends in Microbiology* **2022,** *30* (8), 722-735.
644 37. L. F. Huergo; G. Chandra; M. Merrick, P_{II} signal transduction proteins: nitrogen regulation and
645 beyond. *Fems Microbiology Reviews* **2013,** *37* (2), 251-283.
646 38. Müller, M.-C.; Wagner, T., The Oxoglutarate Binding Site and Regulatory Mechanism Are
647 Conserved in Ammonium Transporter Inhibitors GlnKs from *Methanococcales*. *International Journal of*
648 *Molecular Sciences* **2021,** *22* (16).
649 39. Dodsworth, J. A.; Cady, N. C.; Leigh, J. A., 2-Oxoglutarate and the PII homologues $NifI_1$ and
650 $NifI_2$ regulate nitrogenase activity in cell extracts of *Methanococcus maripaludis*. *Mol Microbiol* **2005,** *56*
651 (6), 1527-38.
652 40. Maslać, N.; Sidhu, C.; Teeling, H.; Wagner, T., Comparative Transcriptomics Sheds Light on
653 Remodeling of Gene Expression during Diazotrophy in the Thermophilic Methanogen
654 *Methanothermococcus thermolithotrophicus*. *Mbio* **2022,** *13* (6).
655 41. Prasse, D.; Schmitz, R. A., Small RNAs Involved in Regulation of Nitrogen Metabolism. *Microbiol*
656 *Spectr* **2018,** *6* (4).
657 42. Wedler, F. C.; Shreve, D. S.; Kenney, R. M.; Ashour, A. E.; Carfi, J.; Rhee, S. G., Two
658 Glutamine Synthetases from *Bacillus caldolyticus*, an Extreme Thermophile. ISOLATION,
659 PHYSICOCHEMICAL AND KINETIC PROPERTIES. *J Biol Chem* **1980,** *255* (19), 9507-16.
660 43. Wang, Q.; Min, C.; Yan, T. T.; Pu, H. F.; Xin, Y. Q.; Zhang, S. Q.; Luo, L.; Yin, Z. M.,
661 Production of glutamine synthetase in *Escherichia coli* using SUMO fusion partner and application to l-
662 glutamine synthesis. *World J Microb Biot* **2011,** *27* (11), 2603-2610.
663 44. Jespersen, M.; Pierik, A. J.; Wagner, T., Structures of the sulfite detoxifying F_{420}-dependent
664 enzyme from *Methanococcales*. *Nat Chem Biol* **2023**.
665 45. Jespersen, M.; Wagner, T., How a methanogen assimilates sulfate: Structural and functional
666 elucidation of the complete sulfate-reduction pathway. *bioRxiv* **2022,** *2022.10.18.512691*.
667 46. Kurth, J. M.; Müller, M.-C.; Welte, C. U.; Wagner, T., Structural Insights into the Methane-
668 Generating Enzyme from a Methoxydotrophic Methanogen Reveal a Restrained Gallery of Post-
669 Translational Modifications. *Microorganisms* **2021,** *9* (4).
670 47. Lemaire, O. N.; Müller, M. C.; Kahnt, J.; Wagner, T., Structural Rearrangements of a
671 Dodecameric Ketol-Acid Reductoisomerase Isolated from a Marine Thermophilic Methanogen.
672 *Biomolecules* **2021,** *11* (11).
673 48. Lemaire, O. N.; Infossi, P.; Ali Chaouche, A.; Espinosa, L.; Leimkuhler, S.; Giudici-Orticoni,
674 M. T.; Méjean, V.; Iobbi-Nivol, C., Small membranous proteins of the TorE/NapE family, crutches for
675 cognate respiratory systems in Proteobacteria. *Sci Rep* **2018,** *8* (1), 13576.
676 49. Vonrhein, C.; Flensburg, C.; Keller, P.; Sharff, A.; Smart, O.; Paciorek, W.; Womack, T.;
677 Bricogne, G., Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallogr D* **2011,** *67*,
678 293-302.
679 50. Tickle, I. J.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G.,
680 STARANISO. http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi. *Global Phasing Ltd., Cambridge,*
681 *United Kingdom*.
682 51. Kabsch, W., *XDS*. *Acta Crystallographica Section D-Biological Crystallography* **2010,** *66*, 125-
683 132.
684 52. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan,
685 R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N.

686 S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., Overview of the *CCP4* suite
687 and current developments. *Acta Crystallogr D* **2011,** *67*, 235-242.
688 53. Liebschner, D.; Afonine, P. V.; Baker, M. L.; Bunkóczi, G.; Chen, V. B.; Croll, T. I.; Hintze,
689 B.; Hung, L. W.; Jain, S.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Poon, B. K.; Prisant, M. G.;
690 Read, R. J.; Richardson, J. S.; Richardson, D. C.; Sammito, M. D.; Sobolev, O. V.; Stockwell, D. H.;
691 Terwilliger, T. C.; Urzhumtsev, A. G.; Videau, L. L.; Williams, C. J.; Adams, P. D., Macromolecular
692 structure determination using X-rays, neutrons and electrons: recent developments in *Phenix*. *Acta*
693 *Crystallogr D* **2019,** *75*, 861-877.
694 54. Bricogne, G.; Blanc, E.; Brandl, M.; Flensburg, C.; Keller, P.; Paciorek, W.; Roversi, P.;
695 Sharff, A.; Smart, O. S.; Vonrhein, C., Buster Version 2.10.4. *Cambridge, United Kingdom: Global*
696 *Phasing Ltd.* **2017**.
697 55. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and development of *Coot*. *Acta*
698 *Crystallographica Section D-Biological Crystallography* **2010,** *66*, 486-501.
699 56. Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.;
700 Murray, L. W.; Richardson, J. S.; Richardson, D. C., *MolProbity*: all-atom structure validation for
701 macromolecular crystallography. *Acta Crystallogr D* **2010,** *66*, 12-21.
702 57. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W. Z.; Lopez, R.; McWilliam,
703 H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G., Fast, scalable generation of high-quality
704 protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **2011,** *7*.
705 58. Robert, X.; Gouet, P., Deciphering key features in protein structures with the new ENDscript
706 server. *Nucleic Acids Research* **2014,** *42* (W1), W320-W324.
707 59. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J. H.; Zhang, Z.; Miller, W.; Lipman, D.
708 J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic*
709 *Acids Research* **1997,** *25* (17), 3389-3402.
710 60. Kumar, S.; Stecher, G.; Tamura, K., MEGA7: Molecular Evolutionary Genetics Analysis Version
711 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **2016,** *33* (7), 1870-1874.
712 61. Jones, D. T.; Taylor, W. R.; Thornton, J. M., *The rapid generation of mutation data matrices*
713 *from protein sequences*. *Comput Appl Biosci* **1992,** *8* (3), 275-282.
714 62. Schoch, C. L.; Ciufo, S.; Domrachev, M.; Hotton, C. L.; Kannan, S.; Khovanskaya, R.; Leipe,
715 D.; Mcveigh, R.; O'Neill, K.; Robbertse, B.; Sharma, S.; Soussov, V.; Sullivan, J. P.; Sun, L.; Turner,
716 S.; Karsch-Mizrachi, I., NCBI Taxonomy: a comprehensive update on curation, resources and tools.
717 *Database-Oxford* **2020**.
718 63. Letunic, I.; Bork, P., Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display
719 and annotation. *Bioinformatics* **2007,** *23* (1), 127-128.
720 64. Crooks, G. E.; Hon, G.; Chandonia, J. M.; Brenner, S. E., WebLogo: A Sequence Logo
721 Generator. *Genome Res* **2004,** *14* (6), 1188-1190.

722 **Figures and Tables**
723
724





E

Enzyme	Substrate	AppK_m (µM)	AppV_{max} (U.mg^{-1})	Hill coefficient
*Mt*GS	NH_4Cl	144.2 ± 12.7	32.7 ± 1.1	1.5 ± 0.2
	Glutamate	8,754.7 ± 865.1	34.2 ± 1.3	1.7 ± 0.2
*Ms*GS	NH_4Cl	72.5 ± 7.0	1.0 ± 0.0	1
	Glutamate	32,570.4 ± 3,133.2	1.4 ± 0.1	2.4 ± 0.4

725
726
727
728 **Figure 1: Purification and characterization of GS from methanogenic archaea.** (A) SDS-PAGE and
729 (B) Native PAGE of 2 µg purified *Mt*GS and *Ms*GS. (C) Relative activity of *Mt*GS at different 2OG
730 concentrations. (D) The specific activity of *Ms*GS with or without 2 mM 2OG at 40 mM glutamate and 20
731 mM NH_4Cl. (E) Kinetic parameters of both GS for NH_4Cl and glutamate. All activities were measured in
732 triplicates.

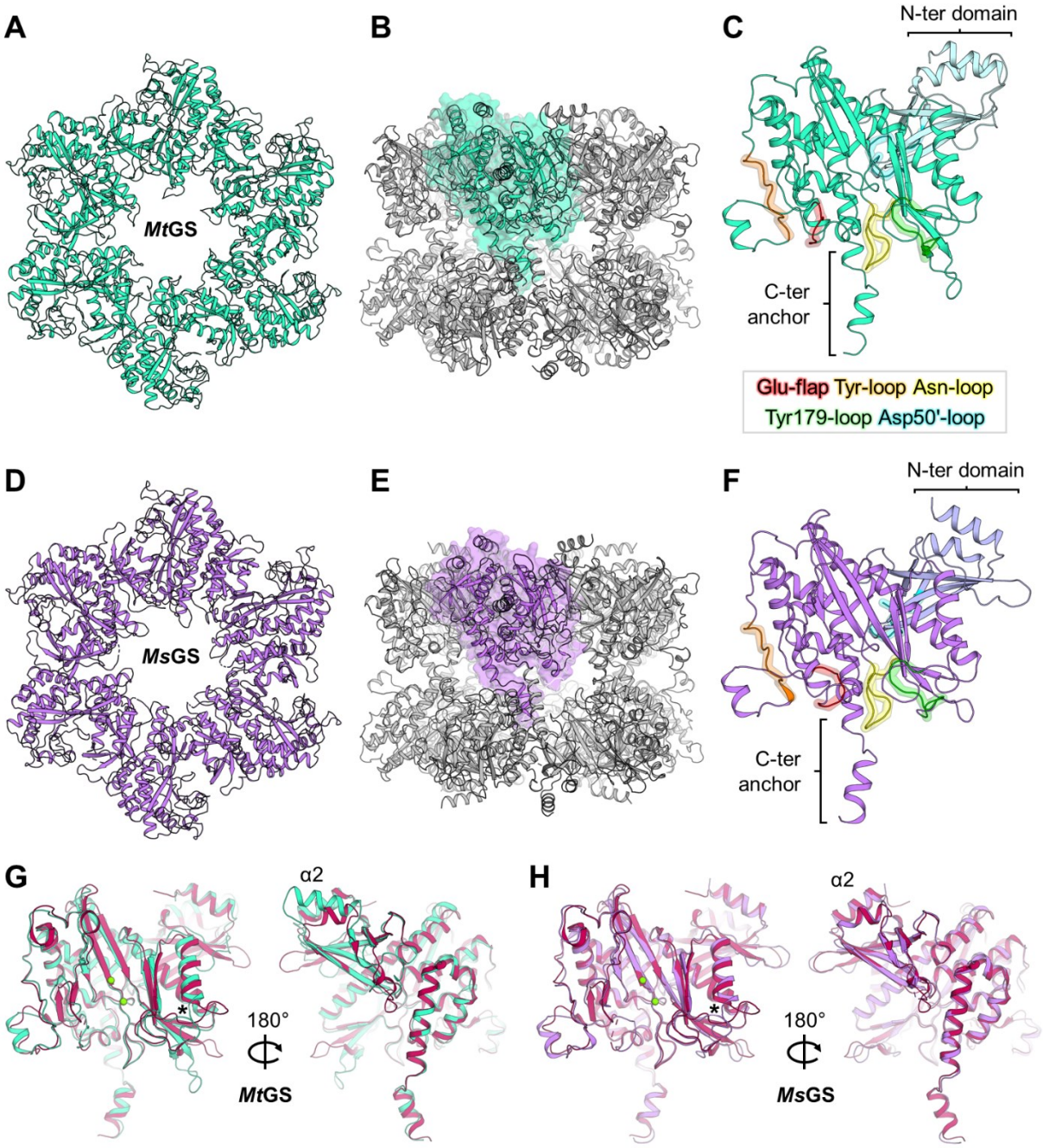


Figure 2: Structural organization of archaeal GS. All models are represented in cartoons. (A) Top view of one *Mt*GS hexamer. (B) Side view of the *Mt*GS dodecamer with one subunit shown in a transparent surface (cyan). (C) *Mt*GS monomer with the main loops highlighted. The color coding of the loops is indicated in the box. The N- and C- terminal domains are colored light blue and cyan, respectively. (D) Top view of one *Ms*GS hexamer. (E) The side view of the *Ms*GS dodecamer with one subunit is shown as a transparent surface (purple). (F) *Ms*GS monomer with the main loops highlighted with the same color coding as in panel C. The N-and C- terminal domains are colored lavender and purple, respectively. (G) Overlay of *Mt*GS (cyan) and *Bs*GS apo state (red, PDB 4LNN). (H) Overlay of *Ms*GS (purple) and *Bs*GS apo state (red). For panels G and H, the star indicates the position of the loop deviating in both archaeal GS compared to *Bs*GS. Mg atoms are displayed as green spheres.

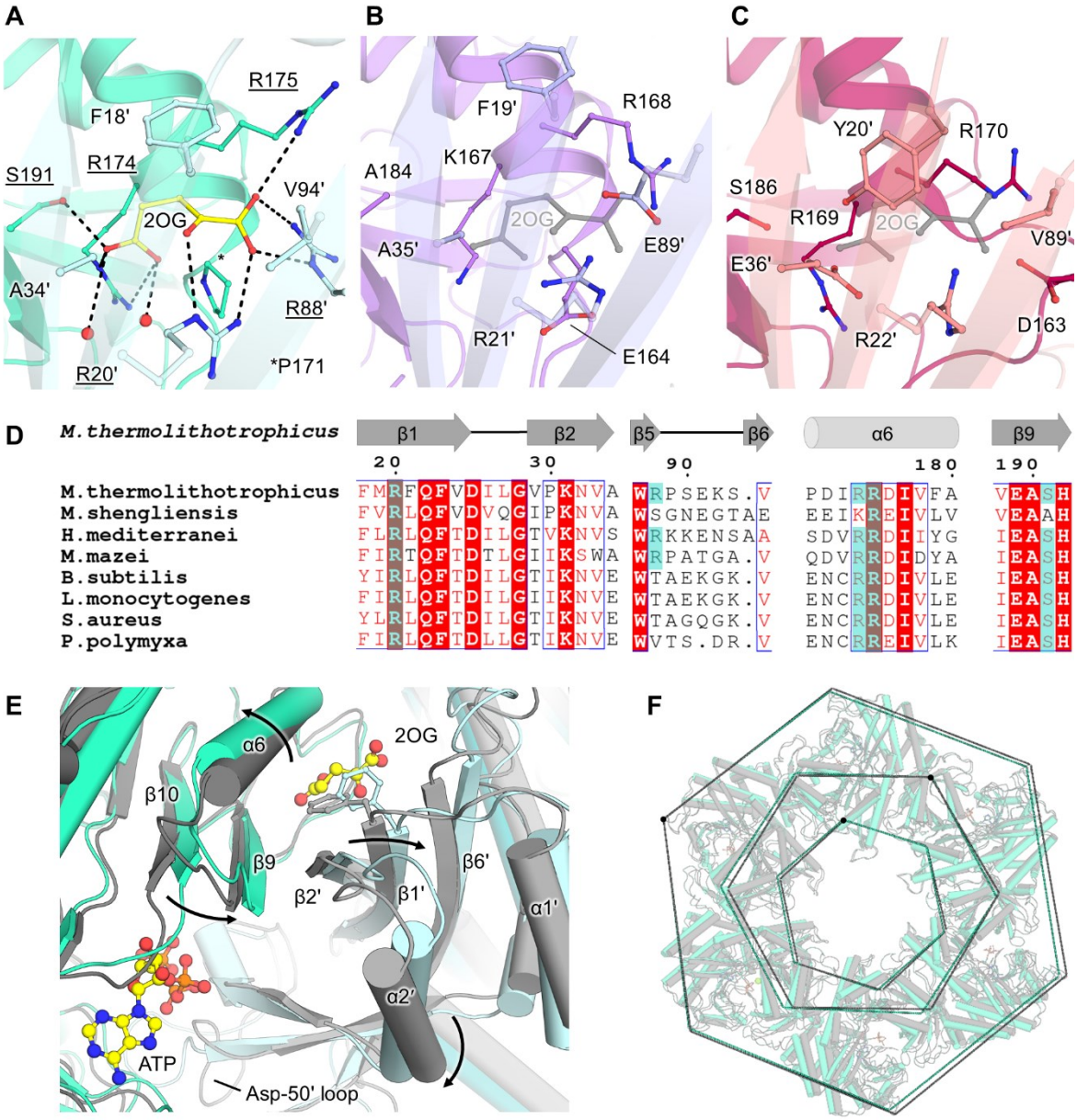


Figure 3: 2OG binding site and structural rearrangement in *Mt*GS. (A) Close-up of the 2OG binding site in *Mt*GS (cyan cartoon, the adjacent monomer in light blue). 2OG and the residues in its vicinity are shown as balls and sticks with contacts in black dashes. (B-C) Same view as in A showing *Ms*GS apo (B, purple cartoon, with the adjacent subunit in light purple) and *Bs*GS (C, red cartoon, with the adjacent subunit in light red, PDB 4LNN). 2OG from *Mt*GS (grey) was superposed to visualize the clash with E89' for *Ms*GS and E36'/Y20' for *Bs*GS. (D) Sequence alignment of different GSI-α in which 2OG-binding residues observed in *Mt*GS are highlighted with a cyan box (see Fig. S7 for the entire alignment). (E) Structural rearrangements between the apo (grey cartoon) and 2OG/Mg^{2+}/ATP bound state (cyan cartoon). The adjacent monomer is colored lighter. Phe18' is shown as sticks. Arrows highlight the movements caused by 2OG binding. (F) *Mt*GS apo (grey cartoon) superposed to *Mt*GS-2OG/Mg^{2+}/ATP (cyan cartoon). The superposition was done on one monomer, and a dashed line was drawn on the Cα position of Val4, Gly198, and Asn267 to illustrate the overall movements. For all, oxygen, nitrogen, and phosphorus are colored in red, blue, and orange, respectively. Carbons are colored depending on the chain and in yellow for ligands.

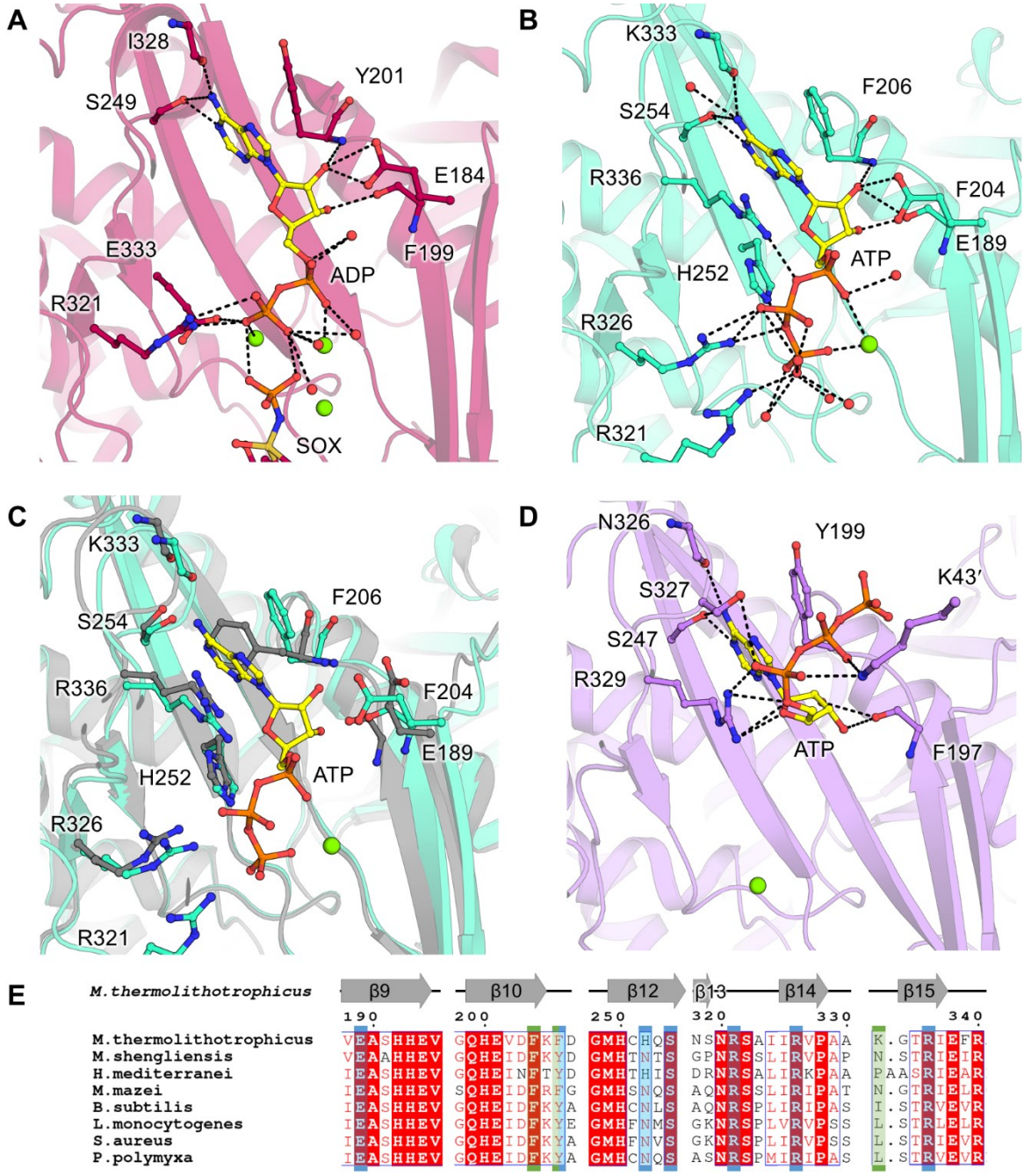


Figure 4: ATP-binding site comparison between different GSI-α. (A) ATP binding site in *Bs*GS transition state (containing SOX/Mg^{2+}/ADP, PDB 4LNI) and (B) *Mt*GS-2OG/Mg^{2+}/ATP. (C) Superposition of the C-terminal domain of *Mt*GS apo (grey) on *Mt*GS-2OG/Mg^{2+}/ATP (cyan), with an overlay of the ATP-binding residues. (D) ATP binding site in *Ms*GS-Mg^{2+}/ATP. Models are represented in transparent cartoons with the ligands (yellow) and interacting residues shown as balls and sticks. Oxygen, nitrogen, sulfur, phosphorus, and magnesium are colored red, blue, dark yellow, orange, and green, respectively. Carbons are colored by chain and ATP carbons in yellow. Hydrogen bonds are visualized as black dashes. (E) Sequence alignment of the ATP binding residues. Residues coordinating the nucleotide via side chain and main chain hydrogen bonds are highlighted by a blue and green box, respectively (based on *Mt*GS). *Ms*GS K43'/S327 were omitted from the analysis due to the artefactual γ-phosphate position.

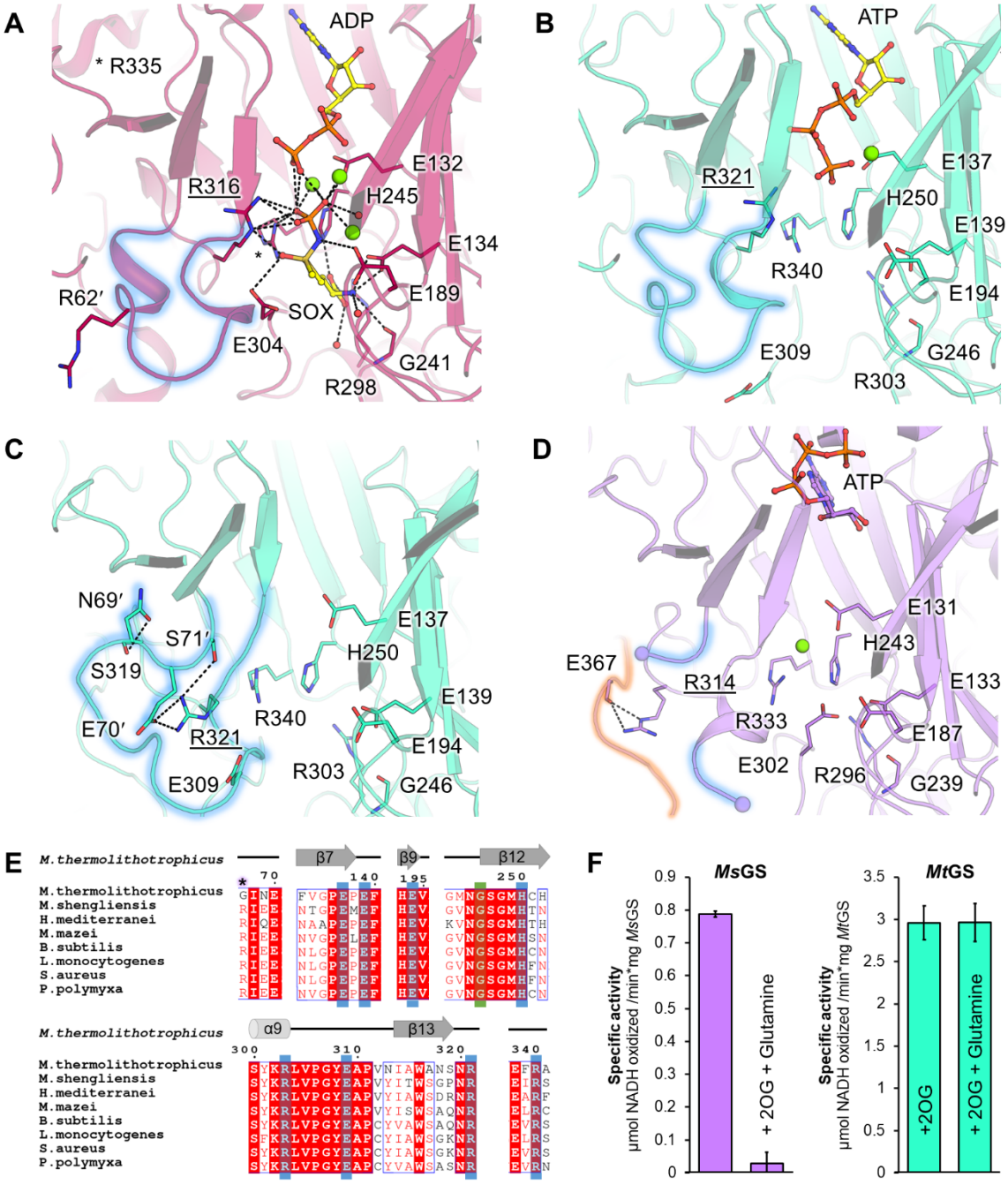


Figure 5. Glutamate-binding site comparison and glutamine feedback inhibition. (A-D) Glutamate binding site in (A) *Bs*GS-SOX/Mg^{2+}/ADP (red, 4LNI), (B) *Mt*GS-2OG/Mg^{2+}/ATP (cyan), (C) *Mt*GS apo (cyan), and (D) *Ms*GS-Mg^{2+}/ATP (purple). Models are in cartoons with ligands and equivalent residues binding SOX as balls and sticks. Oxygen, nitrogen, sulfur, phosphorus, and magnesium are colored red, blue, dark yellow, orange, and green, respectively. Carbons are colored by chain and ATP carbons in yellow. Hydrogen bonds are visualized as black dashes. The Asp-50' loop region is highlighted by a blue glow. For *Ms*GS, the Tyr-loop is highlighted with an orange glow. (E) Alignment of the residues involved in glutamate binding (based on *Bs*GS). Side chain and main chain interactions are highlighted by a blue and green box, respectively. The arginine responsible for glutamine feedback inhibition in *Bs*GS is highlighted by a star. (F) Specific activity in the absence and presence of glutamine in both archaeal GS.

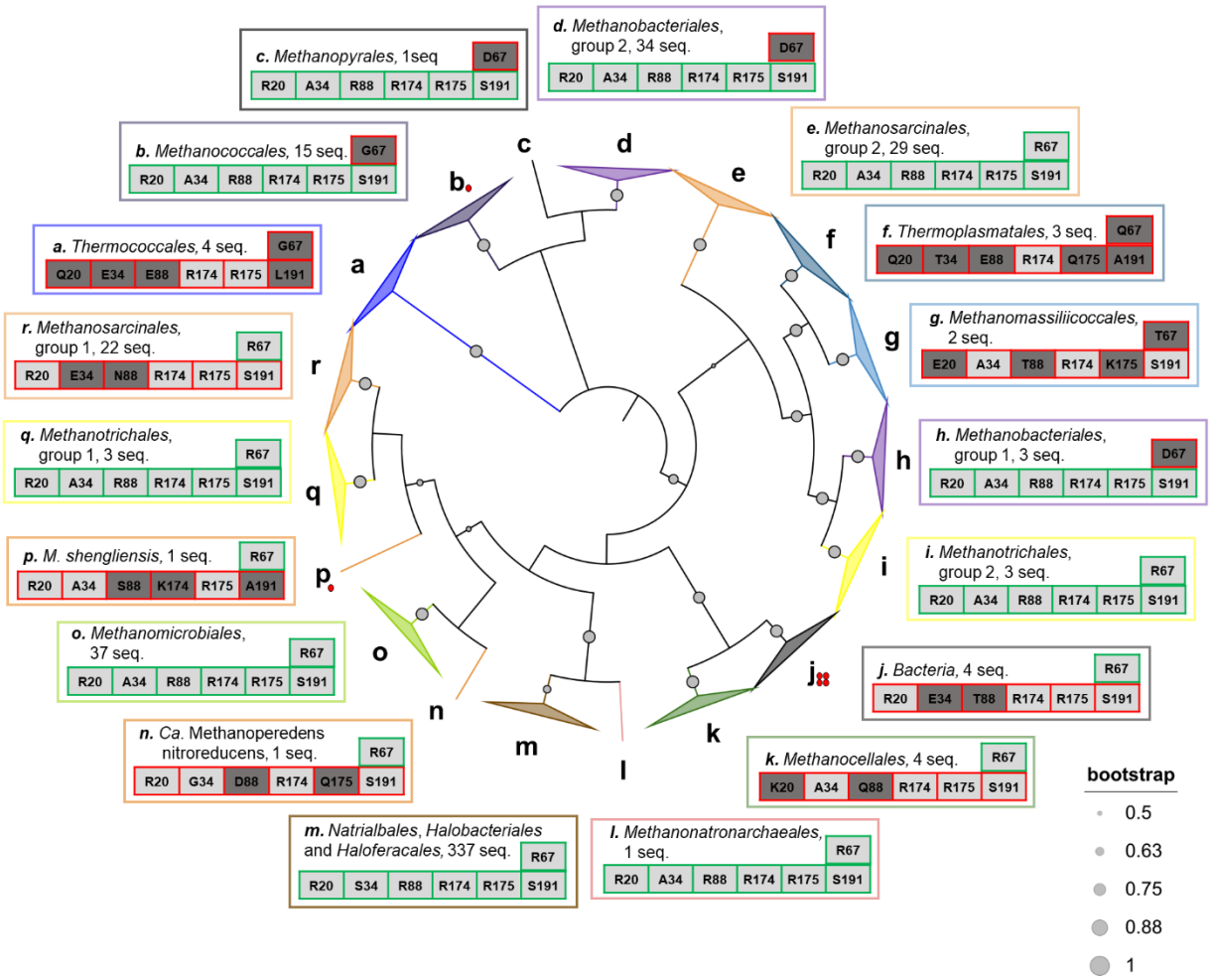


Figure 6. Conservation of residues binding 2OG and glutamine in archaeal GS. The presented phylogenic tree (maximum likelihood) was constructed with the 500 closest sequences to *Mt*GS in the RefSeq database, restricted to the domain *Archaea*, as well as the sequences of the bacterial GSI-α from *B. subtilis*, *S. aureus*, *L. monocytogenes* and *P. polymyxa*. The tree was colored by orders (except for *Bacteria*), and grey dots with different radii represent the bootstrap support of each branch. Branches containing monophyletic groups are collapsed. The most common residues at positions involved in 2OG (Arg20, Arg88, Arg174, Arg175 and Ser191, *Mt*GS numbering, bottom line) and glutamine (position 67 in *Mt*GS, upper line) according to multiple sequence alignment are presented. Ala34 in *Mt*GS is not involved in 2OG coordination, but a substitution by a bulky residue (e.g. glutamate) would hinder its fixation. The residues are colored in light grey or dark grey depending on whether they allow metabolite binding or not, respectively. The residue distribution for each position in each group is presented in Fig. S15, and the sequences used in Table S3. The GS that are predicted to be able or unable to bind 2OG and glutamine are framed in green and red, respectively. Red dots indicate the GSI-α structurally characterized previously or in the present work. The branch containing 337 sequences belonging to *Natrialbales*, *Halobacteriales* and *Haloferales* orders is condensed as no clear monophyletic groups could have been extracted.