

Article

Tertiary and Quaternary Structure Organization in GMP Synthetases: Implications for Catalysis

Lionel Ballut ¹, Sébastien Violot ¹, Frédéric Galisson ¹, Isabelle R. Gonçalves ², Juliette Martin ¹, Santosh Shivakumaraswamy ³, Loïc Carrique ¹, Hemalatha Balam ³ and Nushin Aghajari ^{1,*}

- ¹ Molecular Microbiology and Structural Biochemistry, UMR5086 CNRS-University of Lyon1, 7 Passage du Vercors, CEDEX 07, F-69367 Lyon, France; lionel.ballut@ibcp.fr (L.B.); sebastien.violot@ibcp.fr (S.V.); frederic.galisson@ibcp.fr (F.G.); juliette.martin@ibcp.fr (J.M.); loic.carrique@strubi.ox.ac.uk (L.C.)
- ² Microbiologie, Adaptation et Pathogénie, UMR5240 CNRS-University of Lyon1, F-69622 Villeurbanne, France; isabelle.goncalves@univ-lyon1.fr
- ³ Molecular Biology and Genetics Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore 560064, India; porustg@gmail.com (S.S.); hb@jncasr.ac.in (H.B.)
- * Correspondence: nushin.aghajari@ibcp.fr; Tel.: +33-04-72722633

Abstract: Glutamine amidotransferases, enzymes that transfer nitrogen from Gln to various cellular metabolites, are modular, with the amidotransferase (GATase) domain hydrolyzing Gln, generating ammonia and the acceptor domain catalyzing the addition of nitrogen onto its cognate substrate. GMP synthetase (GMPS), an enzyme in the de novo purine nucleotide biosynthetic pathway, is a glutamine amidotransferase that catalyzes the synthesis of GMP from XMP. The reaction involves activation of XMP through adenylation by ATP in the ATP pyrophosphatase (ATPPase) active site, followed by channeling and attack of NH₃ generated in the GATase pocket. This complex chemistry entails co-ordination of activity across the active sites, allosteric activation of the GATase domain to modulate Gln hydrolysis and channeling of ammonia from the GATase to the acceptor active site. Functional GMPS dimers associate through the dimerization domain. The crystal structure of the Gln-bound complex of *Plasmodium falciparum* GMPS (*Pf*GMPS) for the first time revealed large-scale domain rotation to be associated with catalysis and leading to the juxtaposition of two otherwise spatially distal cysteinyl (C113/C337) residues. In this manuscript, we report on an unusual structural variation in the crystal structure of the C89A/C113A *Pf*GMPS double mutant, wherein a larger degree of domain rotation has led to the dissociation of the dimeric structure. Furthermore, we report a hitherto overlooked signature motif tightly related to catalysis.

Keywords: glutamine amidotransferase; GMP synthetase; *Plasmodium falciparum*; conformational changes; ammonia channel; allosteric regulation; signature motifs; dimeric interface; crystal structure; phylogenetic analysis



Citation: Ballut, L.; Violot, S.; Galisson, F.; Gonçalves, I.R.; Martin, J.; Shivakumaraswamy, S.; Carrique, L.; Balam, H.; Aghajari, N. Tertiary and Quaternary Structure Organization in GMP Synthetases: Implications for Catalysis. *Biomolecules* **2022**, *12*, 871. <https://doi.org/10.3390/biom12070871>

Academic Editor: Beatrice Golinelli-Pimpaneau

Received: 11 May 2022

Accepted: 20 June 2022

Published: 23 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Purine metabolites are essential players in cellular processes and constitute one of the most abundant classes of molecules within mammalian cells. Besides being essential for DNA and RNA synthesis, they play a crucial role in providing cellular energy and intracellular signaling [1]. The synthesis of these highly important molecules, which are involved in energy transfer and metabolic regulation in both eukaryotes and prokaryotes, can follow two routes: either a de novo or a salvage pathway. Within this context, members of the glutamine amidotransferase enzyme family play a crucial role in the biosynthesis of nucleotides and amino acids. In humans, guanosine monophosphate synthase (GMP synthetase and/or GMPS) is required for the de novo biosynthesis of GMP and has, among others, been reported to play a major role in the invasion and tumorigenicity of cells from some human metastatic melanomas, with increased levels of GMPS in these latter compared to what is observed for primary melanomas [2]. Moreover, very recent results suggest that

GMPS may be involved in the progression of cervical cancers, where it may be a marker of unfavorable prognosis of this cancer form [3]. In insects, GMPS has been documented to have a double function in that, besides being required for de novo GMP synthesis, it is involved in transcription control [4]. As concerns parasites for which the purine salvage pathway constitutes the only source of purine nucleotides, the enzyme has been shown to be essential for the viability and infectivity of the parasite *Trypanosoma brucei* [5]. For GMPS' present in the de novo synthesis pathway in plants, the enzyme is used as a target for the design of herbicides [6], while their homologue in the encapsulated yeast *Cryptococcus neoformans* has been reported to be essential for virulence since, in its absence, the yeast cells become guanine auxotroph, leading to compromise in the production of key virulence factors [7]. Finally, in bacteria, inactivation of the GMPS (*guaA*) in *Clostridioides difficile* strain 630 led to cell death in minimal growth conditions, but not in a rich medium. Importantly, the capacity of a *guaA* mutant to colonize the mouse gut was significantly reduced. Together, these results demonstrate the importance of de novo GMP biosynthesis in *C. difficile* during infection [8].

GMP synthetase (EC 6.3.5.2) is a glutamine amidotransferase that catalyzes the final step of de novo GMP biosynthesis, whereby it converts xanthosine 5'-monophosphate (XMP) to guanosine 5'-monophosphate (GMP). GMPS is made up of two domains, one being a glutaminase (GATase) domain that catalyzes the hydrolysis of the amino acid glutamine to glutamate and ammonia. The generated ammonia is used by the ATP pyrophosphatase (ATPPase) domain to catalyze the amination of XMP to form the end product, GMP. It should be noted that whereas in eukaryotes and bacteria, these two domains are located on the same polypeptide chain, the catalytic reactions are carried out by two independent polypeptides in archaea. Hence, GMPS' found in the first category are referred to as two-domain GMPS', while those from the second category are named two-subunit GMPS'. Among the two-domain enzymes, substantial differences have been observed between the GMPS from the parasite *Plasmodium falciparum*, PfGMPS, and the human homolog, thereby establishing these enzymes as attractive drug targets. An immediate example of these differences is that while the human enzyme has been reported to be a functional monomer [9], that of the parasite, together with other parasitic protozoa and prokaryotes, functions as dimers.

In an attempt to contribute to the detailed understanding of the mechanistic aspects of GMPS', we have undertaken a series of in-depth studies of PfGMPS to establish the relations between the structure, function and activity of these enzymes [10,11]. We, among others, have shown that upon binding of Gln to the GATase domain, this latter undergoes an 85° rotation and that this conformational change is required for channeling ammonia from this domain to the ATPase domain in *Plasmodium falciparum* GMP synthetase. Upon this dramatic large-angle rotation, Cys113 (GATase domain) and Cys377 (ATPPase domain) are brought together, engendering the formation of a disulfide bridge in the crystal, as seen in the three-dimensional structure of PfGMPS_C89A (mutation of one of the catalytic residues in the GATase domain) in complex with Gln. Moreover, we confirmed that the obligate functional dimer is formed via the ATPase domain, as also observed in GMPS structures from prokaryotic organisms: *Escherichia coli* (PDB-ID 1GPM, [12]); *Thermus thermophilus* (2YWB, unpublished); *Coxiella burnetti* (3TQI, [13]); *Neisseria gonorrhoeae* (5TW7, unpublished); *Acinetobacter baumannii* (7SBC, unpublished).

Compared to their prokaryotic and parasitic counterparts, mammalian GMP synthetases display a 130 amino acid insert, and at least the *Homo sapiens* GMPS has been reported to be a functional monomer [9]. In order to get further insights into the functioning that distinguishes these glutamine amidotransferases, we performed detailed comparative studies of the existing dimer interfaces, but also compared these latter with the corresponding area in the human GMPS. Furthermore, we performed a phylogenetic analysis of GMP synthetases from various kingdoms of life to obtain a more fine-tuned understanding of the relevance of the extra-domain as found in mammalian GMPS' compared to the homologues that lack this latter, including a number of eukaryotes. A lid-loop, essential for the key

steps of the catalytic reaction and present in the ATPase catalytic site, was also studied and compared with the homologous stretch in GMPS' from other species. Finally, we have performed structural studies of the double mutant *Pf*GMPS_C89A_C113A, which sheds further light on domain rotation and its role in the functioning of the enzyme.

2. Materials and Methods

2.1. Cloning of *Pf*GMPS and Mutagenesis

The three proteins studied herein were fused to an N-terminal 6His tag. The mutant *Pf*GMPS_C89A was constructed using pETP*f*GMPS [14] as a template. Mutant C113A was constructed by the single primer method by introducing a restriction site at the codon to be mutated, followed by knockout of the site to generate the desired mutation [10]. Native *P. falciparum* GMPS, *Pf*GMPSwt, was constructed as described earlier [14] from a native gene from PF10_0123 available in the *P. falciparum* genome database PlasmoDB (<http://www.plasmodb.org> (accessed on 1 January 2007)). The plasmid pETP*f*GMPS served as a template for the C113A mutation. Phusion DNA polymerase from Thermo Scientific was employed for the PCRs, and the resulting PCR products were digested with DpnI from New England Biolabs prior to transformation. The clones were verified by DNA sequencing.

2.2. Protein Purification

Protein expression and purification of the three enzymes studied herein were carried out as described earlier [10,14]. Briefly, *Pf*GMPSwt was expressed from the pQE30*Pf*GMPS construct in *E. coli*, and the pETP*f*GMPS mutant constructs were transformed into BL21-CodonPlus(DE3)-RIL. Hereafter, colonies were inoculated into 5 mL Terrific Broth containing $100 \text{ mg} \times \text{mL}^{-1}$ ampicillin and $34 \text{ mg} \times \text{mL}^{-1}$ chloramphenicol and grown overnight at 37 °C. Inoculum (0.5%) of the overnight culture was then added to 50 mL Terrific Broth supplemented with antibiotics, grown at 37 °C for 6 H and thereafter used at 0.5% to inoculate 2.8 L of Terrific Broth supplemented with antibiotics. Induction with 0.5 mM IPTG (isopropyl- β -D-thiogalactopyranoside) was performed after culture growth to an OD₆₀₀ of 0.6, and the growth continued for 18 H at 18 °C. Following harvesting and storage (−80 °C), the cell pellet was resuspended in 50 mM Tris-HCl, pH 7.4, 10% (v/v) glycerol, 0.1 mM DTT (dithiothreitol), and 0.1 mM PMSF (phenylmethanesulfonylfluoride), which will be referred to as Buffer A. The cell suspension was supplemented with 1 mM PMSF and sonicated on ice, and the cell lysate was centrifuged ($30,500 \times g$ for 45 min). The supernatant of this latter was then loaded onto 1 mL HisTrap HP columns (GE Healthcare) using an ÄKTA HPLC (GE Healthcare). The column was washed with Buffer A containing 1 M NaCl, followed by step gradients of 20 to 100 mM imidazole in Buffer A. Proteins were eluted with Buffer A containing 1 M imidazole. Protein fractions were pooled and concentrated employing 30 kDa cut-off Amicon Ultra-15 centrifugal filter units from EMD Millipore. A final step of size-exclusion chromatography using a HiLoad 16/600 Superdex 200 pg column (GE Healthcare) was carried out in 20 mM Tris-HCl, pH 7.4, 10% (v/v) glycerol, 1 mM EDTA and 2 mM DTT, and the aliquots of the purified protein were flash frozen and stored at −80 °C. The yield was 5 mg of pure protein from 1 liter of culture.

2.3. Crystallization, Data Collection and Structure Determination

Crystallization condition screening was carried out at 292 K (vapor-diffusion in sitting drops), using commercially available crystallization kits. For screening, a Mosquito[®] crystallization robot from STP Labtech was employed using two protein/crystallization agent ratios (200 nL + 200 nL and 300 nL + 100 nL drops equilibrated against 70 μL in MRC Crystallization Plates (Molecular Dimensions)). The mutant protein *Pf*GMPS_C89A_C113A was concentrated to $13 \text{ mg} \times \text{mL}^{-1}$ in 50 mM Tris-HCl pH 8.0, 100 mM NaCl buffer. Crystals grew in a mixture of 16% polyethylene glycol (PEG) 3350, 0.06 M sodium citrate (pH 2.3) and 0.04 M Bis-Tris propane (pH 9.7). Prior to data collection, the crystal was

briefly soaked in a precipitant solution to which 15% ethylene glycol (final concentration) had been added and flash frozen in liquid nitrogen.

X-ray diffraction data were collected on a single crystal (ID30A-3—ESRF, Grenoble, France) at a wavelength of 0.87290 Å. Data processing was done using programs from the XDS package [15]. Phases and experimental electron density maps were calculated after a molecular replacement search using the *PfGMPSwt* three-dimensional structures [10] of the C89A single mutant in complex with glutamine (PDB-ID 4WIO), the native structure of the isolated GATase domain (PDB-ID 4WIN) and the ATPase domain extracted from the full-length structure of the native enzyme (PDB-ID 4WIM) as search models with the PHASER [16] program. The initial model was built using Phenix AutoBuild [17]. Hereafter, alternating cycles of model building and refinement were done employing Coot [18] and REFMAC [19]/Phenix [20].

2.4. Small Angle X-ray Scattering

For SAXS data collection, the samples were concentrated to $\sim 12 \text{ mg} \times \text{mL}^{-1}$. Data were collected on beamline BM29 (ESRF, Grenoble, France) at a wavelength of 0.9919 Å. The data were analyzed using the ATSAS program package [21]. Ab initio models (50 in total) of *PfGMPS_C89A_C113A* were generated from the experimental data. After averaging the 50 generated models, they were filtered to generate the final model.

2.5. Characterization of GATase Domain Rotation

Two dynamic domains, corresponding to the GATase and ATPase structural domains, were identified by the program DynDom [22] from pdb files of *PfGMPSwt*, *PfGMPS_C89A/Gln* and *PfGMPS_C89A_C113A* (monomer of each structure). Hereafter, UCSF ChimeraX v1.3 [23] was used to superimpose *PfGMPSwt* or *PfGMPS_C89A/Gln* onto *PfGMPS_C89A_C113A*, holding the ATPase domain fixed. After superposing the ATPase domains, the change in orientation of the GATase domains was established using ChimeraX [23], which gave the rotation angle and translation along the rotation axis.

2.6. Phylogenetic Data Analysis

Homologs were searched against the UniProtKB database using the program BLASTP [24] and *EcGMPS* and *HsGMPS* as query sequences (Supplementary Table S1). To ensure the correct representation of the different taxonomic groups, a total of 42 BLASTP searches were run on the NCBI website. Herein, each search set was limited to one of the taxonomic groups, as represented in Supplementary Table S1. All sequences that aligned at least 80% of the length of the two and/or three canonical domains (GATase, ATPase, extra-domain) of *EcGMPS/HsGMPS* sequences were kept for multiple sequence alignment. Hereafter, sequences were clustered with CD-HIT to maintain a significant number of sequences for each group. To further enrich the phylogenetic tree, sequences were manually added: 20 well-known model organisms taken from the Uniprot database, 11 from prokaryotes [25], and the sequences corresponding to the existing three-dimensional structures in the protein databank at Rutgers [26]. Multiple sequence alignment was done with the MAFFT program [27] and conserved blocks were selected employing BMGE 1.12 [28]. The BLOSUM62 [29] matrix was used in both programs. 408, 158 and 252 sites were kept for the full-length protein, the GAT domain and the ATPase domain, respectively, for further analysis after character trimming performed by BMGE [28]. Phylogenetic analyses were performed with the LG model and gamma correction, employing a maximum-likelihood approach with PhyML [30]. 100 bootstrap replicates were performed. The phylogenetic tree was generated and rendered with iTOL software [31].

2.7. Bioinformatic Analysis

For the 90 selected sequences used for phylogeny, sequence analyses were done using the Bioinformatics Toolkit resource at <https://toolkit.tuebingen.mpg.de> (accessed on 4 April 2022) [32,33]. Clustal Omega v1.2.1 [34] was employed for alignment of the protein

sequences, and secondary structure predictions of aligned sequences were performed by PSIPRED [35] employing the Ali2D Tool of the Bioinformatics Toolkit resource. Sequence logos were displayed using the WebLogo 3 resource (<http://weblogo.threeplusone.com> (accessed on 15 April 2022)) [36].

2.8. Figure Rendering

Figures with three-dimensional structure representations were rendered using PyMOL [37] and UCSF ChimeraX v1.3 [23].

3. Results

Although the first structural details of a GMPS were published in 1994 [12] and since then 12 more structures of full-length GMPS' have been added to the pdb databank, some aspects of the molecular mechanism of catalysis by this enzyme still remain elusive. In our previous studies of a *Pf*GMPS mutant in which one of the catalytic residues in the GATase domain, Cys89, was mutated to an alanine, the three-dimensional structure in complex with the substrate Gln revealed that this domain was rotated by 85° as compared to the wild-type enzyme [10]. Here, we present novel findings that add further insight into the functioning of this enzyme with a structure of a variant in which not only the catalytic Cys89 was mutated, but also Cys113, which was earlier shown to form a disulfide bridge with Cys377 in the *Pf*GMPS_C89A/Gln complex.

3.1. Crystal- and Solution Structures

The *Pf*GMPS_C89A/C113A mutant enzyme was crystallized, and the crystals of this double mutant belonged to the monoclinic space group C2. The crystal structure of this mutant solved by molecular replacement was determined to a 2.8 Å resolution. The asymmetric unit contains 1 molecule, and all residues have been modeled with the exception of flexible regions with no electron density (residues 234–237, 342–343, 390–396, 487–495 and 521–538). The data collection and refinement statistics are given in Table 1.

Table 1. Data collection and refinement statistics.

	<i>Pf</i> GMPS_C89A_C113A
Data collection	
Beamline	ID30-A ESRF
Wavelength (Å)	0.87290
Resolution range (Å)	30.0–2.80
Space group	C2
Cell dimensions	
a, b, c (Å)	111.65, 64.54, 56.33
α, β, γ (°)	90.00, 90.03, 90.00
number of reflections	24,244
number of unique reflections	9228
R _{meas} (%)	11.8 (46.5)
CC $\frac{1}{2}$	0.99 (0.74)
I/ σ (I)	9.3 (3.6)
Multiplicity	2.6 (2.7)
Completeness (%)	92.3 (93.5)
number of molecules/asymm. unit	1
Refinement	
R/R _{free} ¹ (%)	27.4/31.4
number of atoms	
protein	4070
water	10
Average B-factor (Å ²)	
protein	48.7
water	21.5

Table 1. Cont.

		<i>Pf</i> GMPS_C89A_C113A
RMSD		
	Bond lengths (Å)	0.005
	Angles (°)	1.005
Ramachandran		
	Favored (%)	76.3
	Allowed (%)	18.5
	Outliers (%)	5.2

¹ R_{free}, R_{factor} calculated from 5% of the data excluded from refinement.

When comparing the crystal structure of this double mutant with the crystal structures of the native enzyme and with that of the C89A/Gln mutant complex, substantial conformational changes were observed. The GATase domain of the double mutant (*Pf*GMPS_C89A_C113A) is rotated by ~170° and translated ~1.5 Å compared to the native structure and when holding the ATPase domain fixed. If compared with the structure of the single mutant, the GATase domain is rotated by ~120° and translated ~4 Å, again when holding the ATPase domain fixed (Figure 1).

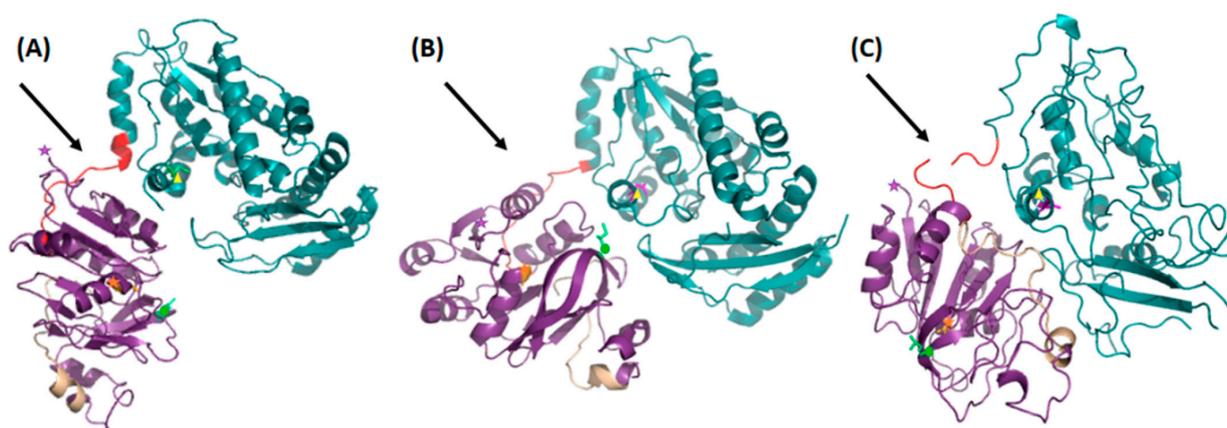


Figure 1. Crystal structures of (A) *Pf*GMPSwt, (B) *Pf*GMPS_C89A in complex with Gln and (C) *Pf*GMPS_C89A_C113A. GATase domains are colored in purple, the linker regions in red, the ATPase domains in teal and the unique (parasite enzyme) insertion including residues 14–159 in wheat. In order to visualize the changes in orientation between the domains with the ATPase domain held fixed, the N-terminus has been labeled with a purple star, C89 (catalytic residue in the GATase domain) with an orange star, C113 (residue forming a disulfide bridge with C377 as observed in the crystal structure of *Pf*GMPS_C89A in complex with Gln) with a green dot and D371 (catalytic residue in the ATPase domain) with a yellow triangle, respectively. When comparing the three ATPase domains, it was noticed that this latter is completely destructured in the double mutant described herein (C). The linker is positioned in front of the GATase domain in the native enzyme (A), behind the GATase domain in the single mutant enzyme (B) and is unstructured in the double mutant enzyme (C).

The electron density of this double mutant revealed a crystal structure that is not only more collapsed than the wild-type and single mutant enzymes described earlier but also significantly less structured, resulting in bad geometry, as can be observed from the Ramachandran parameters. Destructuring is particularly important in the ATPase domain, and to our surprise, the observed structure of this double mutant is a monomeric enzyme. Moreover, the collapsed nature of the structure has resulted in many contacts being very close. Here, it should be mentioned that numerous attempts to grow better crystals of this mutant failed.

The monomeric nature of the double mutant in solution was confirmed by Small Angle X-ray Scattering studies (Figure 2), as opposed to the solution dimers observed earlier for the *Pf*GMPS_{wt} and *Pf*GMPS_C89A enzymes [10] and confirmed herein.

Indeed, *Pf*GMPS_C89A_C113A lacks Gln-dependent GMPS activity due to the mutation of the catalytic cysteine and shows a 6.7-fold decrease in K_m for NH_4Cl compared with the wild-type enzyme and a 2.9-fold decrease compared with the single mutant *Pf*GMPS_C89A. With k_{cat} values being almost similar, *Pf*GMPS_C89A_C113A had the highest k_{cat}/K_m value for ammonia-dependent GMP formation. Nevertheless, the conservation of enzyme activity shows that this double mutant may also adopt a dimeric form. However, the structure of *Pf*GMPS_C89A_C113A shows that the double mutant at the extreme rotation of the GATase domain transitions from a dimer to a monomer conformation. Careful inspection of the three-dimensional structure of the *Pf*GMPS_C89A_C113A mutant suggests that this destructuring is a result of an aberrant rotation of the GATase domain in conjunction with a unique insertion (Asp141-Ile159; Figure 1) in this same domain being specific to the parasite enzyme [38]. Taken together, these issues engender the destabilization of the dimer interface, which has been shown to be essential for enzyme activity. Unequivocally, the double mutant is no longer able to form this disulfide bridge, resulting in an enzyme that, in principle, can continue its rotation. Here, it should be noted that Cys377 is only found in *Plasmodium* and Cys113 only in *P. falciparum*, *P. vivax* and *P. knowlesi*. Interestingly, Cys377 is part of the so-called lid-loop, which precedes the helix that holds Asp371 and Glu374 (catalytic residues of the ATPase domain) and is, as evident from our earlier mutagenesis studies, a key residue in the AMP-XMP intermediate formation. The formation of a disulfide bridge in this position has only been observed in the *Pf*GMPS_C89A/Gln crystal structure among the GMPS structures solved to date.

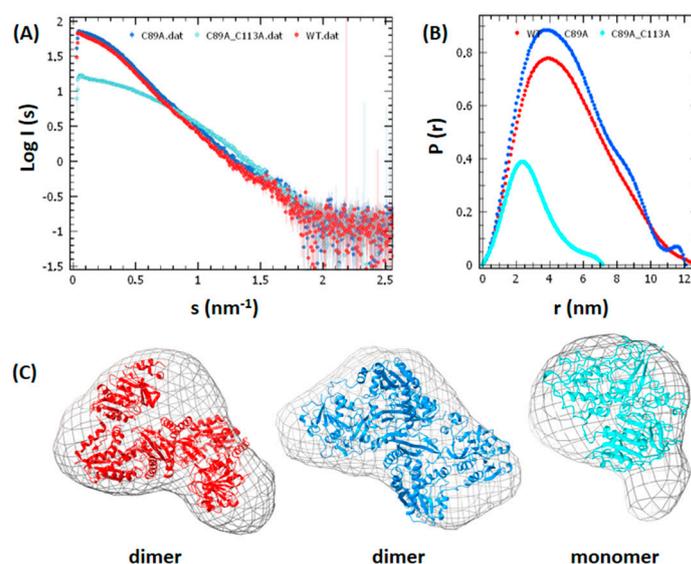


Figure 2. Solution structures of *Pf*GMPS_{wt}, *Pf*GMPS_C89A and *Pf*GMPS_C89A_C113A as determined by Small Angle X-ray Scattering. The (A) radially scattered X-ray intensity (I) as a function of the scattering angle (s), with *Pf*GMPS_{wt} dimer (red), *Pf*GMPS_C89A dimer (blue) and *Pf*GMPS_C89A_C113A monomer (cyan). (B) Pair-distance distribution functions $P(r)$ with the same color codes as for (A). (C) Crystal structures with the same color coding as above the respective variants superimposed onto the averaged ab initio shape reconstructed with DAMMIF [39].

It should be noted that to measure the rotation of the GATase domain in solution, a FRET-based assay was conceived. For this experiment, Trp167 from the GATase domain was chosen as the FRET donor, and Cys377 from the ATPase domain was covalently labeled with 5-(2-((2-iodoacetyl)amino)ethylamino)naphthalene-1-sulfonic acid “IAEDANS” as the FRET (Förster Resonance Energy Transfer) acceptor. Inspection of the crystal structures of native *Pf*GMPS_{wt} and the rotated *Pf*GMPS_C89A indicates that upon the 85° rotation of

the GATase domain, the distance between these two residues is reduced by approximately 14 Å, from ~19.5 Å to ~5.5 Å. To ensure that only the above-mentioned Trp and Cys were labeled, the second tryptophan and the 10 remaining cysteines in the enzyme were mutated to phenylalanine and alanines, respectively. Unfortunately, this resulted in a mutated enzyme that was insoluble.

Altogether, one may conclude that whereas this double mutant (*PfGMPS_C89A_C113A*) can adopt a dimeric form, as concluded from the conservation of the enzyme activity, it is also able to transit to a monomer conformation at the extreme rotation of the GATase domain, as confirmed by the crystal structure and Small Angle X-ray Scattering studies.

3.2. Comparative Studies of Dimerization Interfaces

Considering that all our recent (described above) and earlier [10,11] observations of *PfGMPSwt* indicate that in two-domain dimer GMPS', the dimer interface is essential for forming a scaffold allowing interaction with the substrate XMP, we decided to elaborate on this issue. A close-up on the GMPS dimer interfaces, allowing a detailed inspection of the interacting residues in this area (Figure 3 and Supplementary Figure S1), indicates that these interactions are conferred by the C-terminus of the enzyme. Comparison of the dimer interfaces of GMPS' of known structure reveals that these interactions are partly highly conserved among the different species (Figure 3) and involve two arginines, two glutamic acids, an aspartic acid and an aromatic residue from one monomer and an arginine from a second monomer. Some of these residues (Glu553, Phe554 and Glu555, *PfGMPS* numbering) are involved in the highly conserved C-terminal signature motif KPPXTXE(F/W)X. Within this motif, Glu553 has been determined to be critical for XMP binding (mutation to Leu), whereas for Glu555 the Gln and NH₄Cl dependent GMP formation by the *PfGMPS_E555L* mutant was only marginally lower than what was observed for the wild-type enzyme [11].

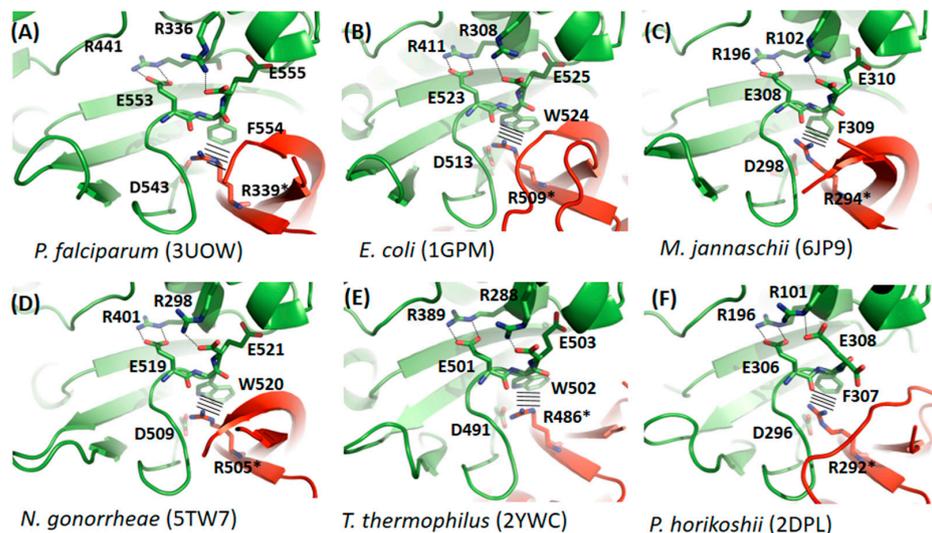


Figure 3. Conservation of amino acid residues involved in interactions with XMP in the ATPase domain and the stabilization of the dimer interface. Lines indicate π —cation interactions. For GMPS from (A) *P. falciparum*, (B) *E. coli*, (C) *M. jannaschii*, (D) *N. gonorrhoeae*, (E) *T. thermophilus* and (F) *P. horikoshii* residues from monomer A are shown in green and monomer B in red with the interacting arginines marked by an asterisk.

A second part of the interface is less conserved (Supplementary Figure S1A–F) and rather seems to be specific for the organism as *eg. T. thermophilus* GMPS which is stabilized by a salt bridge, which is a frequent stabilizing factor in proteins from thermophilic species.

For mammalian GMPS', as exemplified by *HsGMPS*, which is a monomeric enzyme, the overall fold differs from that seen for functional dimeric enzymes as *eg. PfGMPS*, in that *HsGMPS* displays an extra-domain (known as the D1 sub-domain) made up of 130 amino

acid residues. This subdomain, which is present in mammalian GMPS', is involved in substrate binding but also in dimerization, as seen in the crystal structure of *HsGMPS*, which turned out to be a dimer in the crystal [40]. This extra-domain insertion is located in the ATPase domain (Figure 4), and the orientation of the D1 subdomain and the subsequent C-terminal part (referred to as subdomain D2 by Welin and colleagues [40]), mimics the dimer interface as formed by ATPase domains of two monomers seen in functional dimers such as that of the *P. falciparum* parasite (Supplementary Figure S2). Moreover, as also indicated by the authors, when comparing subdomains D1 and D2 within the *HsGMPS*, they display a similar polypeptide fold when the two sub-structures are superimposed. It should be highlighted that the aforementioned "subsequent" C-terminal part is present in all GMP synthetases (Figures 3 and 4) and includes, among others, the highly conserved C-terminal signature motif, as mentioned above. As seen in Supplementary Figure S2 and Figure 3, the residues involved in the interactions with the XMP substrate are intriguingly highly conserved, despite the fact that part of the residues stem from the extra-domain in the human enzyme instead of a second molecule, as is the case for the dimeric GMPS'. To summarize, when comparing the interactions seen in GMPS' of known three-dimensional structures, in one part of the region corresponding to the dimer interface in dimers as exemplified by *PfGMPS* and the region mimicking this same region in monomers as *eg. HsGMPS* conservation of amino acid residues involved in interactions with XMP and in the stabilization of the dimer interface are observed. As opposed hereto, the second part of the interface seems more specific to the organism and shows a lower degree of conservation.

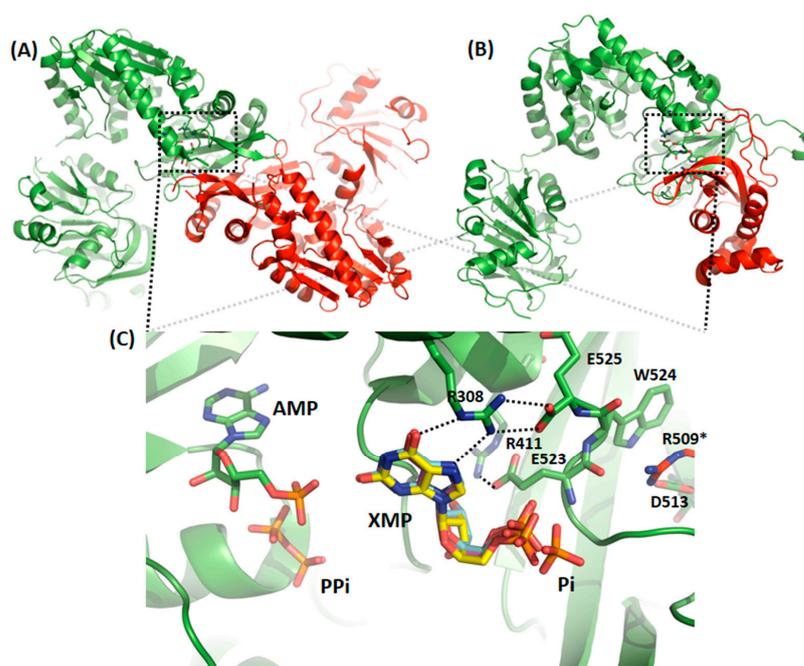


Figure 4. Overall three-dimensional structure of GMP Synthetases from a two-domain dimeric enzyme as exemplified by (A) *P. falciparum* GMPS dimer with one monomer colored in green and the other in red, versus that of a two-domain monomeric enzyme as exemplified by the (B) *H. sapiens* GMPS monomer in which the extra-domain is colored in red. The rectangle in the stippled lines delineates the substrate-binding site in the ATPase domain in the two structures. (C) Close-up on the ATPase active site in structures with bound substrate (XMP in *HsGMPS*—PDB 2VXO, yellow; in *TiGMPS*—PDB 2YWC, light blue; in *PfGMPS*—PDB 3UOW, pink and AMP, PP_i and P_i in *EcGMPS*, green and orange). To simplify, only protein residues (in green) of the *E. coli* enzyme are shown in this figure. As can be seen from the interaction scheme, arginine 308 (see numbering correspondence for the other GMPS' in Figure 3) interacts with the C-terminal and together with the remaining residues allow the interaction with XMP and the stabilization of the interface corresponding to the so-called dimerization interface.

3.3. Phylogenetic Analyses

Based on these considerations and in order to get a better understanding of the evolution of these multi-domain architectural proteins, for which two different quaternary structures constitute the physiologically active forms, we decided to look at the phylogenetic relationships between these enzymes.

As a means to start with representative samples of these two groups, the GMPS protein sequence from *E. coli* and/or *H. sapiens* was used in a NCBI blastp search against 42 different proteomes belonging to Bacteria, Archaea and Eukaryotes. The retrieved sequences were then clustered to maintain a representative number of sequences for each group. Nineteen sequences of model organisms were then added along with sequences of bacteria used in an earlier phylogenetic analysis [25]. Different trees were constructed either using the complete protein sequences or by limiting the analysis to a single domain (i.e., GATase or ATPase) and similar results were obtained.

For each tree, a clear separation occurred between a group (ensemble A) of eukaryotic proteins and a group (ensemble B) of both prokaryotic and eukaryotic proteins (Figure 5 and Supplementary Figures S3 and S4). Eukaryotic organisms found in ensemble A belong to the Opisthokonta, Apusozoa, Amoebozoa, Euglenozoa, Haptophyta and Cryptophyta. Ensemble A gathers protein sequences displaying (in the ATPase domain) a conserved “extra-sequence” (residues 450-578 in *H. sapiens* GMPS corresponding to the dimerization domain D1) absent in the protein sequences found in ensemble B.

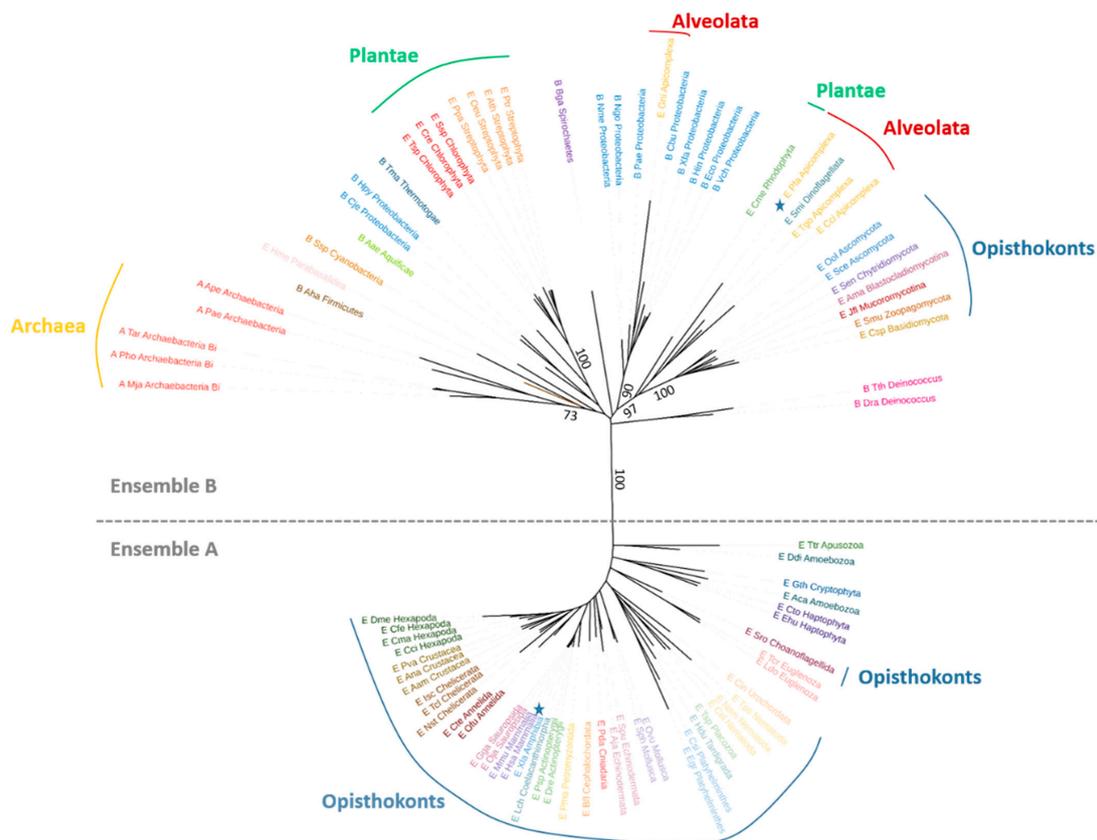


Figure 5. Phylogenetic tree of full-length GMPS enzymes. A ML phylogeny based on 408 amino acid alignment positions of 87 sequences was constructed with PhyML [30] using GMPS’ from eukaryotes, archaea and bacteria. *PfGMPS* and *HsGMPS* have been labeled with stars. A clear separation between a group of eukaryotic proteins (ensemble A) and a group of both prokaryotic and eukaryotic proteins (ensemble B) can be observed.

Based on our above-standing comparative analysis of the three-dimensional structures of GMPS’ available in the Protein Data Bank RCSB [26], we observed that the site mimicking

the dimer interface is both issued from the extra-sequence and another part of the protein sequence, which is present in all GMP Synthetases (Figure 4). Prediction of the expected secondary structures of this extra-domain for each of the studied sequences from ensemble A indicates conserved β -strands and α -helices (Figure S5). In accordance with the monophyly of ensemble A proteins in the phylogeny, the complex structural rearrangements, which end up by the acquisition of the extra-domain existing in these proteins, suggest that they have certainly occurred only once during evolution in the ancestor of the ensemble A proteins.

In ensemble B, the extra-sequence is systematically absent, suggesting that the active form of all these proteins displays a dimeric organization. Organisms for which the GMPS three-dimensional structures have been solved and which form dimers (*E. coli*, 1GPM; *T. thermophilus*, 2YWC; *P. horikoshii*, 2DPL; *C. burnetti*, 3TQI; *P. falciparum* 3UOW and 4WIM; *N. gonorrhoea*, 5TW7; *M. jannashii*, 6JP9; *A. baumannii*, 7SBC) are all present in this second group supporting this hypothesis. Ensemble B is constituted by bacteria, archae and eukaryotes. Among the eukaryotes, we find Plantae (Streptophyta, Chlorophyta and Rhodophyta), Fungi, Parabasalidea and Alveolata (Dinoflagellata and Apicomplexa). No species possessing either type of protein was found.

The presence of the extra-domain in most Opisthokonts, but also in Apusozoan and Amoebozoa GMP Synthetases, suggests that it was present in the ancestral Opisthokonta proteins. Surprisingly, among the Opisthokonta, fungal GMPS' do not possess the extra domain. Fungal sequences form a monophyletic group in ensemble B, suggesting the replacement of an ancestral Opisthokonta version of the gene (with the extra-sequence) by a laterally transferred gene. Such a GMPS-encoding lateral gene transfer (LGT) was suggested in an earlier study [25].

3.4. Comparative Studies of the Catalytic Loop Region in the ATPase Domain

Upon rotation of the GATase domain, Cys113 is brought in close proximity to Cys377, as exemplified in the crystal structure of PfGMPS_C89A/Gln (Figure 1). This latter cysteine takes part in the loop (lid-loop, residues 376–401) ensuing the helix, which holds the catalytic Asp371, a key residue in the AMP-XMP intermediate formation. This same helix also holds the catalytic Glu374 (invariant across all species), which, when mutated to Leu, results in an enzyme in which inter-domain cross-talk upon ATP- and XMP-binding to the ATPase domain is abolished, and which is thought to be the key residue for allosteric activation of this enzyme. Indeed, these two catalytic residues are believed to participate in the positioning of NH_3 in front of the adenylyl-XMP bond for the attack to form GMP [10]. Once the 85° rotation has taken place, the so-called lid-loop 376-401 is deeply buried into the ATPase domain (Figure 6A) and therefore also contributes, together with the dimer interface, to form a scaffold for the binding of the substrate and ultimately to displace the products of the reaction. The release of the lid-loop most probably triggers the rotation of the GATase domain back to its resting state, a hypothesis supported, among others, by the fact that this latter is not rotated in known crystal structures where the lid-loop is either disordered (all except *HsGMPS*) or ordered but not present in the ATPase pocket, as seen for *HsGMPS* (Figure 6B).

With the aim of studying whether these observations are fortuitous, we performed a bioinformatics analysis to establish the extent to which the catalytic lid-loops are conserved within the two identified phylogenetic ensembles. These studies indicated that within each ensemble, there is an important conservation of both primary and secondary structures (Supplementary Figure S6), whereas they differ across ensembles with the exception of the IK(T/S)HHN motif. Indeed, this latter is essential for the synthesis of the adenylyl-XMP intermediate through residues His388 and His389, as well as the formation of GMP through the interactions of Asn390 [11]. Despite the abovementioned differences, careful inspection reveals that in the vicinity of this loop there are conserved structural elements, such as the α -helix (D331—Y351 (*Pf*)/S332—M355 (*Hs*) in front of the helix bearing the catalytic residues, as well as the helix, which is the first structural element in the ATPase domain

(H241—Y251 (*Pf*)/T218—G235 (*Hs*), Figure 6A,B). While the latter helix does not show particular sequence conservation, the first displays a signature motif, which to the best of our knowledge has not been reported earlier and which corresponds to PEXKRKIIGXXF (Figure 6C,D). Interestingly, these two conserved α -helices appear more or less unwound in the crystal structure of the double mutant (Supplementary Figure S7).

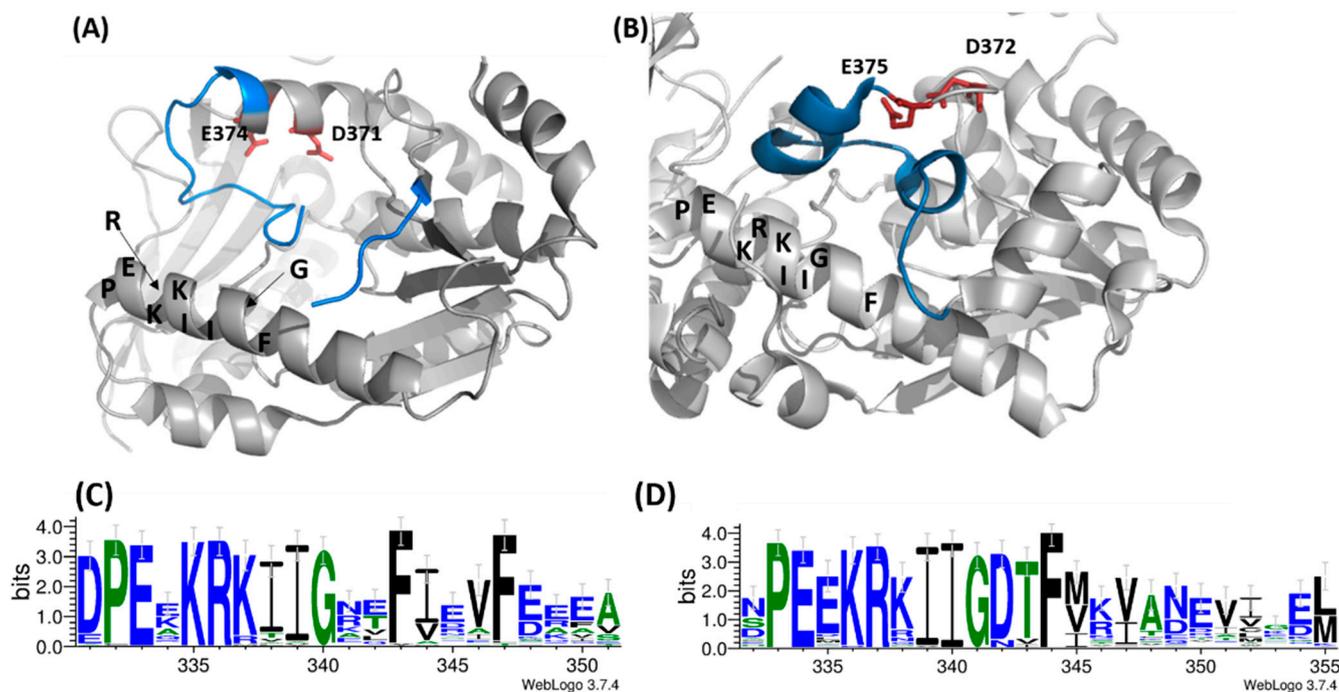


Figure 6. Close-up view of the ATPase substrate binding pocket in (A) *P. falciparum* GMPS and (B) *H. sapiens* GMPS. The catalytic residues Asp371/Asp372 and Glu374/Glu375 in the two structures are shown as orange/red sticks, respectively, and the lid-loop in both structures is colored in blue. It can be observed that, whereas the loop enters the substrate binding pocket in the parasite enzyme, it remains outside the pocket in the mammalian counterpart. Residues of the signature motif PEXKRKIIGXXF on the helix facing the lid-loop are indicated. Weblogo of the same signature motif in (C) dimeric and (D) monomeric GMP synthetases. The selected regions include amino acid residues corresponding to stretch (C) Asp 331—Tyr 351 in the *P. falciparum* enzyme and (D) Ser 331—Met 355 in the human counterpart. The alignments were rendered using WebLogo 3.7.4. [41].

4. Discussion

The three-dimensional structure of the double mutant *Pf*GMPS_C89A_C113A indicated that a rotation of the GATase domain of approximately 170° and a $\sim 1.5 \text{ \AA}$ translation has taken place, as compared to its position in the *Pf*GMPSwt structure. Whether this corresponds to a physiologically relevant form seems unlikely and is in any case uncertain. As opposed hereto, the earlier observed crystal structure of the single mutant *Pf*GMPS_C89A in complex with Gln, in which an 85° rotation had taken place upon the binding of glutamine, seems more realistic. Indeed, rotation of the GATase domain upon addition of glutamine has also been observed in a very recent study of *A. fumigatus* GMPS, for which the rotation was established by mass ion-mobility mass spectrometry, IM-MS [42]. By all means, besides presenting an enzyme with a partly destroyed substrate binding pocket in the ATPase domain (Supplementary Figure S7), in *Pf*GMPS_C89A_C113A, the interactions that are critical for ensuring correct ammonia channeling from the GATase domain to the ATPase domain are modified. It is tempting to think that there is a relation between this observation and the fact that upon rotation, in the native structure, Cys377 may form a disulfide bridge with Cys113, as earlier observed for the C89A mutant structure [10]. Hence, mutation of Cys113 to an alanine precludes the enzyme from resting in

a conformation in which ammonia is in the correct position for attacking the AMP-XMP intermediate, as observed earlier in the 85° rotated enzyme (Supplementary Figure S8). In this sense, C113 could play the role of a barrier/manacle which prevents the enzyme from adopting a monomeric and partly unstructured and inactive form. This makes sense when recalling that C113 is present only in *P. falciparum*, *P. vivax* and *P. knowlesi*, and that the purine nucleotide metabolic pathway constitutes the sole source of purine nucleotides to the rapidly multiplying parasite. As concerns GMPS' from the phylogenic ensemble containing monomeric enzymes, as revealed in this study, this issue is irrelevant since this group of enzymes with their extra-domain mimic the dimer-interface and therefore do not seem to be concerned by the destabilization of the ATPase binding pocket. Indeed, in principle, they should always display the same structure.

The lid-loop, which has been shown to be essential for all the key steps in the catalytic process, including inter-domain communication, enhancement of the substrate affinity, formation of the ammonia channel, expulsion of the reaction products from the substrate binding pocket, etc. [10], appears to be conserved within the two respective ensembles both from a primary and secondary structure point of view. Comparison across the two ensembles, however, indicates that here, only the lid-loop IK(T/S)HHN motif is conserved. Despite this lower degree of conservation when comparing the two ensembles directly, one against the other, it appears that two α -helices are conserved in the immediate surroundings of the lid-loop, and that one of these also holds a signature motif PEXKRKIIGXXF. The tertiary structure conservation of these α -helices throughout the existing three-dimensional structures of GMPS, regardless of their origin, and their unwinding in the inactive double mutant discussed herein, support their importance for the catalytic process. Altogether, our results suggest that future studies should target the residues within the newly discovered signature motif found in one of the two conserved α -helices in order to attribute the role of each of these amino-acid residues in the catalytic reaction.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom12070871/s1>, Figure S1: Stabilization of the dimer interface by amino acid residues that are less conserved across the GMPS' from the studied species; Figure S2: Close-up on the dimer interface; Figure S3: Phylogenetic tree of GATase domains from GMP Synthetases; Figure S4: Phylogenetic tree of ATPase domains from GMP Synthetases; Figure S5: Primary and secondary structure conservation of the extra-domain present in "ensemble A" GMP synthetases; Figure S6: Primary and secondary structure conservation of the catalytic lid-loop in dimeric and monomeric GMP synthetases; Figure S7: Close-up view on the ATPase substrate binding pocket in the *P. falciparum* GMPS_C89A_C113A double mutant; Figure S8: Ammonia channeling in *P. falciparum* GMPS; Table S1: Protein sequences used to generate the phylogenetic tree of PfGMPS homologs.

Author Contributions: Conceptualization, L.B. and N.A.; formal analysis, L.B. and N.A.; investigation, L.B., S.V., F.G., I.R.G., J.M., S.S. and L.C.; writing—original draft preparation, N.A. and L.B.; writing—review and editing, all authors; supervision, L.B. and N.A.; project administration, N.A.; funding acquisition, H.B. and N.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the CNRS, the ANR project grant "PLASMOPUR" ANR-17-CE11-0032, the Fondation Innovation en Infectiologie, "FINOVI project AO12-27" and the Fondation pour la Recherche Médicale (FRM: FDT201604035284). H.B. is a JC Bose National Fellow and acknowledges funding from the Department of Biotechnology and Science and Engineering Research Board, Government of India.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in the RCSB Protein Databank at Rutgers under PDB-ID "7ZU9". The rest of the data presented in this study are available upon reasonable request from the corresponding author.

Acknowledgments: Technical support from staff on beamlines MX and FIP2 (both European Synchrotron Radiation Facility), as well as on beamlines PX3 (Swiss Light Source, Switzerland) and Proxima 1 (Soleil, France) is gratefully acknowledged. We are also grateful for assistance from V. Gueguen-Chaignon, E. Diesis and F. Delolme of the Protein Science Facility of SFR Biosciences Lyon (UAR3444/US8).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pedley, A.M.; Benkovic, S.J. A New View into the Regulation of Purine Metabolism: The Purinosome. *Trends Biochem. Sci.* **2017**, *42*, 141–154. [[CrossRef](#)] [[PubMed](#)]
2. Bianchi-Smiraglia, A.; Wawrzyniak, J.A.; Bagati, A.; Marvin, E.K.; Ackroyd, J.; Moparthy, S.; Bshara, W.; Fink, E.E.; Foley, C.E.; Morozovich, G.E.; et al. Pharmacological Targeting of Guanosine Monophosphate Synthase Suppresses Melanoma Cell Invasion and Tumorigenicity. *Cell Death Differ.* **2015**, *22*, 1858–1864. [[CrossRef](#)] [[PubMed](#)]
3. Wang, J.; Wu, Y.; Li, Y.; Wang, Y.; Shen, F.; Zhou, J.; Chen, Y. Guanosine Monophosphate Synthase Upregulation Mediates Cervical Cancer Progression by Inhibiting the Apoptosis of Cervical Cancer Cells via the Stat3/P53 Pathway. *Int. J. Oncol.* **2021**, *58*, 3. [[CrossRef](#)] [[PubMed](#)]
4. van der Knaap, J.A.; Kozhevnikova, E.; Langenberg, K.; Moshkin, Y.M.; Verrijzer, C.P. Biosynthetic Enzyme GMP Synthetase Cooperates with Ubiquitin-Specific Protease 7 in Transcriptional Regulation of Ecdysteroid Target Genes. *Mol. Cell. Biol.* **2010**, *30*, 736–744. [[CrossRef](#)] [[PubMed](#)]
5. Li, Q.; Leija, C.; Rijo-Ferreira, F.; Chen, J.; Cestari, I.; Stuart, K.; Tu, B.P.; Phillips, M.A. GMP Synthase Is Essential for Viability and Infectivity of Trypanosoma Brucei despite a Redundant Purine Salvage Pathway. *Mol. Microbiol.* **2015**, *97*, 1006–1020. [[CrossRef](#)] [[PubMed](#)]
6. Lerchl, J.; Erhardt, T.; Sonnewald, U.; Boldt, R. GMP Synthetase Derived from Plants 2001. Google Patents AU7658700A, 1 October 2010.
7. Chitty, J.L.; Tatzenko, T.L.; Williams, S.J.; Koh, Y.Q.A.E.; Corfield, E.C.; Butler, M.S.; Robertson, A.A.B.; Cooper, M.A.; Kappler, U.; Kobe, B.; et al. GMP Synthase Is Required for Virulence Factor Production and Infection by *Cryptococcus neoformans*. *J. Biol. Chem.* **2017**, *292*, 3049–3059. [[CrossRef](#)] [[PubMed](#)]
8. Smith-Peter, E.; Séguin, D.L.; St-Pierre, É.; Sekulovic, O.; Jeanneau, S.; Tremblay-Tétreault, C.; Lamontagne, A.-M.; Jacques, P.-É.; Lafontaine, D.A.; Fortier, L.-C. Inactivation of the Riboswitch-Controlled GMP Synthase GaaA in *Clostridioides difficile* Is Associated with Severe Growth Defects and Poor Infectivity in a Mouse Model of Infection. *RNA Biol.* **2021**, *18*, 699–710. [[CrossRef](#)]
9. Lou, L.; Nakamura, J.; Tsing, S.; Nguyen, B.; Chow, J.; Straub, K.; Chan, H.; Barnett, J. High-Level Production from a Baculovirus Expression System and Biochemical Characterization of Human GMP Synthetase. *Protein Expr. Purif.* **1995**, *6*, 487–495. [[CrossRef](#)]
10. Ballut, L.; Violot, S.; Shivakumaraswamy, S.; Thota, L.P.; Sathya, M.; Kunala, J.; Dijkstra, B.W.; Terreux, R.; Haser, R.; Balaran, H.; et al. Active Site Coupling in *Plasmodium falciparum* GMP Synthetase Is Triggered by Domain Rotation. *Nat. Commun.* **2015**, *6*, 8930. [[CrossRef](#)]
11. Shivakumaraswamy, S.; Pandey, N.; Ballut, L.; Violot, S.; Aghajari, N.; Balaran, H. Helices on Interdomain Interface Couple Catalysis in the ATPase Domain with Allostery in *Plasmodium falciparum* GMP Synthetase. *Chembiochem* **2020**, *21*, 2805–2817. [[CrossRef](#)]
12. Tesmer, J.J.; Klem, T.J.; Deras, M.L.; Davisson, V.J.; Smith, J.L. The Crystal Structure of GMP Synthetase Reveals a Novel Catalytic Triad and Is a Structural Paradigm for Two Enzyme Families. *Nat. Struct. Biol.* **1996**, *3*, 74–86. [[CrossRef](#)] [[PubMed](#)]
13. Franklin, M.C.; Cheung, J.; Rudolph, M.J.; Burshteyn, F.; Cassidy, M.; Gary, E.; Hillerich, B.; Yao, Z.-K.; Carlier, P.R.; Totrov, M.; et al. Structural Genomics for Drug Design against the Pathogen *Coxiella burnetii*. *Proteins* **2015**, *83*, 2124–2136. [[CrossRef](#)] [[PubMed](#)]
14. Bhat, J.Y.; Venkatachala, R.; Singh, K.; Gupta, K.; Sarma, S.P.; Balaran, H. Ammonia Channeling in *Plasmodium falciparum* GMP Synthetase: Investigation by NMR Spectroscopy and Biochemical Assays. *Biochemistry* **2011**, *50*, 3346–3356. [[CrossRef](#)] [[PubMed](#)]
15. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [[CrossRef](#)]
16. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [[CrossRef](#)]
17. Terwilliger, T.C.; Grosse-Kunstleve, R.W.; Afonine, P.V.; Moriarty, N.W.; Zwart, P.H.; Hung, L.W.; Read, R.J.; Adams, P.D. Iterative Model Building, Structure Refinement and Density Modification with the PHENIX AutoBuild Wizard. *Acta Crystallogr. D Biol. Crystallogr.* **2008**, *64*, 61–69. [[CrossRef](#)]
18. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [[CrossRef](#)]
19. Vagin, A.A.; Steiner, R.A.; Lebedev, A.A.; Potterton, L.; McNicholas, S.; Long, F.; Murshudov, G.N. REFMAC5 Dictionary: Organization of Prior Chemical Knowledge and Guidelines for Its Use. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2184–2195. [[CrossRef](#)]

20. Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 352–367. [[CrossRef](#)]
21. Manalastas-Cantos, K.; Konarev, P.V.; Hajizadeh, N.R.; Kikhney, A.G.; Petoukhov, M.V.; Molodenskiy, D.S.; Panjkovich, A.; Mertens, H.D.T.; Gruzinov, A.; Borges, C.; et al. ATSAS 3.0: Expanded Functionality and New Tools for Small-Angle Scattering Data Analysis. *J. Appl. Crystallogr.* **2021**, *54*, 343–355. [[CrossRef](#)]
22. Qi, G.; Lee, R.; Hayward, S. A Comprehensive and Non-Redundant Database of Protein Domain Movements. *Bioinformatics* **2005**, *21*, 2832–2838. [[CrossRef](#)] [[PubMed](#)]
23. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure Visualization for Researchers, Educators, and Developers. *Protein Sci.* **2021**, *30*, 70–82. [[CrossRef](#)] [[PubMed](#)]
24. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
25. Yanai, I.; Wolf, Y.I.; Koonin, E.V. Evolution of Gene Fusions: Horizontal Transfer versus Independent Events. *Genome Biol.* **2002**, *3*, 1–13. [[CrossRef](#)]
26. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)] [[PubMed](#)]
27. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
28. Criscuolo, A.; Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): A New Software for Selection of Phylogenetic Informative Regions from Multiple Sequence Alignments. *BMC Evol. Biol.* **2010**, *10*, 210. [[CrossRef](#)]
29. Henikoff, S.; Henikoff, J.G. Amino Acid Substitution Matrices from Protein Blocks. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10915–10919. [[CrossRef](#)]
30. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)]
31. Letunic, I.; Bork, P. Interactive Tree Of Life (ITOL) v4: Recent Updates and New Developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)]
32. Zimmermann, L.; Stephens, A.; Nam, S.-Z.; Rau, D.; Kübler, J.; Lozajic, M.; Gabler, F.; Söding, J.; Lupas, A.N.; Alva, V. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at Its Core. *J. Mol. Biol.* **2018**, *430*, 2237–2243. [[CrossRef](#)] [[PubMed](#)]
33. Gabler, F.; Nam, S.-Z.; Till, S.; Mirdita, M.; Steinegger, M.; Söding, J.; Lupas, A.N.; Alva, V. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. *Curr. Protoc. Bioinform.* **2020**, *72*, e108. [[CrossRef](#)] [[PubMed](#)]
34. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [[CrossRef](#)] [[PubMed](#)]
35. Jones, D.T. Protein Secondary Structure Prediction Based on Position-Specific Scoring Matrices. *J. Mol. Biol.* **1999**, *292*, 195–202. [[CrossRef](#)]
36. Crooks, G.E.; Hon, G.; Chandonia, J.-M.; Brenner, S.E. WebLogo: A Sequence Logo Generator. *Genome Res.* **2004**, *14*, 1188–1190. [[CrossRef](#)]
37. DeLano, W.L. *PyMOL 0.99.Rc6*; DeLano Scientific LLC.: Palo Alto, CA, USA, 2002.
38. Bhat, J.Y.; Shastri, B.G.; Balaram, H. Kinetic and Biochemical Characterization of *Plasmodium falciparum* GMP Synthetase. *Biochem. J.* **2008**, *409*, 263–273. [[CrossRef](#)]
39. Franke, D.; Svergun, D.I. DAMMIF, a Program for Rapid *ab-initio* Shape Determination in Small-Angle Scattering. *J. Appl. Crystallogr.* **2009**, *42*, 342–346. [[CrossRef](#)]
40. Welin, M.; Lehtiö, L.; Johansson, A.; Flodin, S.; Nyman, T.; Trésaugues, L.; Hammarström, M.; Gräslund, S.; Nordlund, P. Substrate Specificity and Oligomerization of Human GMP Synthetase. *J. Mol. Biol.* **2013**, *425*, 4323–4333. [[CrossRef](#)]
41. Schneider, T.D.; Stephens, R.M. Sequence Logos: A New Way to Display Consensus Sequences. *Nucleic Acids Res.* **1990**, *18*, 6097–6100. [[CrossRef](#)]
42. Nguyen, S.; Jovcevski, B.; Pukala, T.L.; Bruning, J.B. Structural Insights into the Antifungal Drug Target Guanosine Monophosphate Synthase from *Aspergillus fumigatus*. *Acta Crystallogr. D Struct. Biol.* **2022**, *78*, 248–259. [[CrossRef](#)]