



**HAL**  
open science

## Biochemical and Structural Characterization of Chi-Class Glutathione Transferases: A Snapshot on the Glutathione Transferase Encoded by sll0067 Gene in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803

Eva Mocchetti, Laura Morette, Guillermo Mulliert, Sandrine Mathiot, Benoît Guillot, Francois Dehez, Franck Chauvat, Corinne Cassier-Chauvat, Céline Brochier-Armanet, Claude Didierjean, et al.

### ► To cite this version:

Eva Mocchetti, Laura Morette, Guillermo Mulliert, Sandrine Mathiot, Benoît Guillot, et al.. Biochemical and Structural Characterization of Chi-Class Glutathione Transferases: A Snapshot on the Glutathione Transferase Encoded by sll0067 Gene in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *Biomolecules*, 2022, 12 (10), pp.1466. 10.3390/biom12101466 . hal-03854391

**HAL Id: hal-03854391**

**<https://hal.science/hal-03854391v1>**





Submitted on 15 Nov 2022

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

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

## Article

# Biochemical and Structural Characterization of Chi-Class Glutathione Transferases: A Snapshot on the Glutathione Transferase Encoded by *sll0067* Gene in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803

Eva Mochetti <sup>1,†</sup>, Laura Morette <sup>2,†</sup>, Guillermo Mulliert <sup>1</sup>, Sandrine Mathiot <sup>1</sup>, Benoît Guillot <sup>1</sup>, François Dehez <sup>3</sup>, Franck Chauvat <sup>4</sup>, Corinne Cassier-Chauvat <sup>4</sup>, Céline Brochier-Armanet <sup>5</sup>, Claude Didierjean <sup>1,\*</sup> and Arnaud Hecker <sup>2,\*</sup>

<sup>1</sup> Université de Lorraine, CNRS, CRM2, F-54000 Nancy, France

<sup>2</sup> Université de Lorraine, INRAE, IAM, F-54000 Nancy, France

<sup>3</sup> Université de Lorraine, CNRS, LPCT, F-54000 Nancy, France

<sup>4</sup> Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), F-91190 Gif-sur-Yvette, France

<sup>5</sup> Université de Lyon 1, CNRS, LBBE, F-69622 Villeurbanne, France

\* Correspondence: claudedidierjean@univ-lorraine.fr (C.D.); arnaud.hecker@univ-lorraine.fr (A.H.)

† These authors contributed equally to this work.



**Citation:** Mochetti, E.; Morette, L.; Mulliert, G.; Mathiot, S.; Guillot, B.; Dehez, F.; Chauvat, F.; Cassier-Chauvat, C.; Brochier-Armanet, C.; Didierjean, C.; et al. Biochemical and Structural Characterization of Chi-Class Glutathione Transferases: A Snapshot on the Glutathione Transferase Encoded by *sll0067* Gene in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *Biomolecules* **2022**, *12*, 1466. <https://doi.org/10.3390/biom12101466>

Academic Editor: Bengt Mannervik

Received: 4 August 2022

Accepted: 5 October 2022

Published: 13 October 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/>).

**Abstract:** Glutathione transferases (GSTs) constitute a widespread superfamily of enzymes notably involved in detoxification processes and/or in specialized metabolism. In the cyanobacterium *Synechocystis* sp. PCC 6803, SynGSTC1, a chi-class GST (GSTC), is thought to participate in the detoxification process of methylglyoxal, a toxic by-product of cellular metabolism. A comparative genomic analysis showed that GSTCs were present in all orders of cyanobacteria with the exception of the basal order Gloeobacterales. These enzymes were also detected in some marine and freshwater noncyanobacterial bacteria, probably as a result of horizontal gene transfer events. GSTCs were shorter of about 30 residues compared to most cytosolic GSTs and had a well-conserved SRAS motif in the active site (<sup>10</sup>SRAS<sup>13</sup> in SynGSTC1). The crystal structure of SynGSTC1 in complex with glutathione adopted the canonical GST fold with a very open active site because the  $\alpha 4$  and  $\alpha 5$  helices were exceptionally short. A transferred multipolar electron-density analysis allowed a fine description of the solved structure. Unexpectedly, Ser10 did not have an electrostatic influence on glutathione as usually observed in serinyl-GSTs. The S10A variant was only slightly less efficient than the wild-type and molecular dynamics simulations suggested that S10 was a stabilizer of the protein backbone rather than an anchor site for glutathione.

**Keywords:** glutathione transferase; glutathione; cyanobacteria; *Synechocystis* sp. PCC 6803; crystallography; biochemistry; phylogeny

## 1. Introduction

Glutathione transferases (GSTs) constitute a widespread superfamily of enzymes playing crucial roles in the cell notably in detoxification processes and in specialized secondary metabolism by catalyzing three major kinds of reactions. These include catalytic reactions where glutathione (GSH) is consumed (GSH-conjugation), reactions where GSH is not consumed (isomerization and dehalogenation) and reactions where GSH is oxidized (thiol-transferase and reduction activities) [1]. At the structural level, canonical GSTs are mainly dimeric proteins, and each subunit adopts a conserved fold composed of an N-terminal thioredoxin (TRX) domain linked to an all  $\alpha$ -helical C-terminal domain. The active site of the enzyme is located in a cleft at the interface between both domains and contains a GSH-binding site (G site) and a hydrophobic-substrate binding site (H site). Depending

on their primary sequence conservation, GSTs were classified into classes designated by a Greek letter. GSTs with a sequence identity greater than 40% belong to the same class, whereas proteins of different classes share less than 25% sequence identity [2]. GSTs were further distinguished into four catalytic types, tyrosine (TyrGSTs), serine (SerGSTs), cysteine (CysGSTs) and atypical (AtyGSTs), depending on an assumed important residue for catalysis [3]. The tyrosine, serine and cysteine residues have a conserved position in the structures. The tyrosine residue is located at the C-terminal end of the first strand ( $\beta$ 1). The serine and cysteine residues have the same position at the N-terminal end of the first helix ( $\alpha$ 1). AtyGSTs do not have a specific residue at a conserved position. In CysGSTs, the cysteine residue has a reactant role according to the enzyme mechanism ontology [4], since it forms a covalent bond with the substrate during the catalytic act. The important residue of the others (TyrGSTs, SerGSTs and AtyGSTs) has a spectator role in enhancing the nucleophilicity of the glutathione thiolate group, and the mutation of this residue does not abolish the activity. Through noncatalytic properties, so-called ligandins, hitherto underestimated compared to the other documented roles of GSTs, many GSTs also participate in the binding and transport of small heterocyclic ligands [5,6].

GSTs have been extensively investigated in animals and plants because of their great relevance to human health and agriculture [7–9]. In contrast, studies in bacteria remain scarce, especially in the cyanobacterial phylum that encompasses oxygenic photosynthetic prokaryotes considered to be the ancestors of chloroplasts. It has been speculated that cyanobacteria may be the first organisms to harbor GSTs [10]. Greek letters have been used for eight classes of GSTs from bacteria: beta, chi, eta, nu, rho, theta, xi and zeta [11,12]. Other classes exist that often have specific functions such as LigE, LigF and LigG involved in lignin degradation in soil bacteria [13,14].

The chi class is thought to be specific of cyanobacteria and three isoforms (TeGSTC1 from *Thermosynechococcus elongatus* BP-1, SeGSTC1 from *Synechococcus elongatus* PCC 6301 and SynGSTC1 from *Synechocystis* sp. PCC 6803) have been characterized biochemically [10,15,16]. A preliminary crystallographic study has been reported for TeGSTC1 and SeGSTC1 [17]. All the three isoforms (TeGSTC1, SeGSTC1 and SynGSTC1) exhibit similar activities. They efficiently catalyze the addition of GSH on various isothiocyanates and show moderate activities toward other classical substrates such as chlorodinitrobenzene [10,15,16]. Interestingly, we recently showed that SynGSTC1 is involved in the detoxification of methylglyoxal, a toxic by-product of the cellular metabolism [15]. Despite a shorter sequence length compared to other GSTs, homology modelling combined with secondary structure prediction suggested that chi GSTs (GSTCs) adopt the fold of canonical GSTs. Their amino acid sequences show two motifs usually found in GSTs. The motif I, which contains an invariant cis-proline residue as well as a  $\beta\beta\alpha$  structure essential for the stabilization of the  $\gamma$ -glutamyl moiety of GSH, is the most conserved region in all of the GSTs [18]. Motif II, in turn, contains a very well conserved aspartic acid important for fold stability [19]. Recently, a conserved tyrosine residue located at the fifth position of the N-terminus of GSTCs has been proposed as the catalytic residue [20]. To better characterize the chi class of GSTs, it was necessary to obtain an experimental three-dimensional model. Therefore, we determined the first crystal structure of a chi-class GST (SynGSTC1), performed a robust phylogenetic study and completed the biochemical data by testing new substrates and modulating the active site residues by site-directed mutagenesis.

## 2. Materials and Methods

### 2.1. Cloning, Mutagenesis, Expression and Purification

SynGSTC1 (SII0067) encoding sequence was amplified by PCR from *Synechocystis* sp. PCC 6803 genomic DNA as template using specific forward and reverse primers containing *Nde*I and *Xho*I restriction sites, respectively (Table S1). The amplified sequence was subsequently digested and cloned in *E. coli* expression vector pET-26b between *Nde*I and *Xho*I restriction sites allowing the fusion of a His-tag at the C-terminal part of SynGSTC1 as previously described [15]. Various catalytic mutants (S10T, S10A, S10C and R11A) were generated

by site-directed mutagenesis using the QuikChange site-directed mutagenesis kit (Agilent Technologies) and specific mutagenic primers listed in Table S1. The sequences have been confirmed by DNA sequencing.

The expression of recombinant SynGSTC1 and variants were performed at 37 °C using *E. coli* Rosetta2 (DE3) pLysS expression strain (Novagen) transformed with appropriate plasmid in LB medium supplemented with kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL). When the cell culture reached an OD<sub>600 nm</sub> of 0.7–0.8, the expression of the SynGSTC1 (or S10T or S10A or S10C or R11A) recombinant protein was induced with 0.1 mM isopropyl β-D-1-thio-galactopyranoside (IPTG) for 4 h at 37 °C. Cells were then harvested by centrifugation, resuspended in a 30 mM Tris-HCl buffer (pH 8.0) supplemented with 200 mM NaCl (lysis buffer) and stored at –20 °C until use. After the lysis of the cells by sonication, the resulting cell extract was centrifuged at 40,000 × g for 20 min at 4 °C to remove cellular debris and aggregated proteins. After the addition of 10 mM imidazole, SynGSTC1 was purified from the soluble extract by gravity-flow chromatography on a nickel nitrilotriacetate (Ni-NTA) agarose resin (Qiagen, Hilden Germany). After a washing step with lysis buffer containing 20 mM imidazole, recombinant SynGSTC1 was eluted using lysis buffer supplemented with 250 mM imidazole. The fractions of interest were pooled, concentrated by ultrafiltration, subjected to a size exclusion chromatography using a Superdex™200 16/600 column connected to an ÄKTA-Purifier™ device (Cytiva) and eluted with lysis buffer. The purified recombinant protein was concentrated and finally stored at –20 °C. The concentration of SynGSTC1 recombinant protein was determined at 280 nm using a theoretical molar absorption coefficient of 28,420 M<sup>-1</sup>·cm<sup>-1</sup>.

## 2.2. Crystallization, X-ray Data Collection, Processing and Refinement

A first screening of 288 crystallization conditions was carried out at the CRM2 crystallogensis platform (University of Lorraine) with an Oryx 8 crystallogensis robot (Douglas Instruments Ltd, Hungerford, UK). Crystals were optimized manually at 4 °C by the microbatch-under-oil method. Solutions of SynGSTC1 and the variants contained 30–40 mg·mL<sup>-1</sup> protein in 30 mM Tris buffer (pH 8.0) supplemented with 200 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine (TCEP) and 10 mM glutathione. SynGSTC1 was crystallized by mixing 1 µL of protein with 1 µL of solution consisting of 16% (*w/v*) PEG 8000, 40 mM potassium phosphate monobasic and 20% (*w/v*) glycerol (condition no. 32, Wizard™ Classic Crystallization Screen III, Rigaku, Tokyo, Japan).

Preliminary X-ray diffraction experiments were carried out in-house on an Agilent SuperNova diffractometer (Rigaku Oxford Diffraction) equipped with a CCD detector. Data collections were carried out at the ESRF, on beamline FIP BM07 (ESRF, Grenoble, France) and (PX1 and PX2, SOLEIL, Gif-Sur-Yvette, France). Data sets were indexed and integrated with XDS [21], and scaled and merged with Aimless [22] from the CCP4 suite [23]. The structure of SynGSTC1 was solved by molecular replacement using *MoRDa* [24] with the coordinates of a GST from *Rhodobacter sphaeroides* (PDB entry 3LSZ) as the search model. Structures of SynGSTC1 and its variants were refined with *BUSTER* [25] and manually improved with *Coot* [26]. The validation of all structures was performed with the PDB validation service (<http://validate.wwpdb.org>, accessed on 30 September 2022). The coordinates and structure factors have been deposited in the Protein Data Bank (PDB entries 8AI8, 8AI9, 8AIB). Crystal data, diffraction and refinement statistics are shown in Table 1.

**Table 1.** Statistics of X-ray diffraction data collection and model refinement.

	Wild-Type	S10T	R11A
<b>Data Collection</b>			
Diffraction source	ESRF-BM07	ESRF-BM07	ESRF-BM07
Detector	Pilatus 6M	Pilatus 6M	Pilatus 6M
Wavelength (Å)	0.97951	0.97951	0.97951

**Table 1.** Cont.

	Wild-Type	S10T	R11A
Space Group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit-cell $a; c$ (Å)	92.5; 193.6	92.9; 193.6	92.2; 193.1
Resolution Range (Å)	48.4 1.7 (1.73 1.70)	48.4 1.7 (1.73 1.70)	46.1 2.2 (2.27 2.20)
Tot. no. of meas. int.	1,200,217 (41158)	1,244,521 (62,785)	503,802 (22,268)
Unique reflections	92,986 (4529)	93,931 (4590)	39,298 (2006)
Average redundancy	13 (9)	13.2 (14)	13 (11)
Mean $I/\sigma(I)$	24.8 (1.8)	17.0 (2.0)	18.4 (2.4)
Completeness (%)	100.0 (99.6)	100.0 (100.0)	91.3 (55.5)
$R_{\text{merge}}$	0.056 (1.097)	0.084 (1.52)	0.097 (1.039)
$R_{\text{meas}}$	0.061 (1.168)	0.087 (1.59)	0.100 (1.142)
$CC_{1/2}$	1.00 (0.83)	1.00 (0.84)	1.00 (0.85)
Wilson $B$ -factor (Å <sup>2</sup> )	29.6	28.6	41.5
<b>Refinement</b>			
Resolution Range (Å)	24.8 1.7	24.2 1.7	31.3 2.2
No. of reflections	92839	93783	39248
$R_{\text{work}}/R_{\text{free}}$	0.204/0.221	0.206/0.221	0.215/0.242
Corr $F_o-F_c/F_o-F_{c_{\text{free}}}$	0.938/0.936	0.940/0.939	0.907/0.888
Total number of atoms	3469	3500	3253
Average $B$ -factor (Å <sup>2</sup> )	34.0	32.5	44.0
<b>Model quality</b>			
RMSZ Bond lengths	0.41	0.42	0.42
RMSZ Bond angles	0.54	0.56	0.56
Ramachandran fav. (%)	98	98	98
Ramachandran all. (%)	2	2	2
Rotamer outliers (%)	0	0	1
Clashscore	1	1	1
PDB entry	8AI8	8AI9	8AIB

$R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - I(hkl)|}{\sum_{hkl} \sum_i I_i(hkl)}$ .  $R_{\text{meas}} = \frac{\sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - I(hkl)|}{\sum_{hkl} \sum_i I_i(hkl)}$ .  $CC_{1/2}$  is the correlation coefficient of the mean intensities between two random half-sets of data.  $R_{\text{work}} = \frac{\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_{hkl} |F_{\text{obs}}|}$ . In total, 5% of reflections were selected for  $R_{\text{free}}$  calculation. RMSZ: root mean square Z-score. The MolProbity clashscore is the number of serious clashes per 1000 atoms. Values in parentheses are for highest resolution shell.

### 2.3. Structure Analysis Based on Electron Density Distribution

To calculate the electrostatic interaction energies between residues of SynGSTC1 active site and the glutathione ligand, the electron charge density of the complex based on the Hansen and Coppens multipolar model [27] was determined (method detailed in Supplementary Materials). The electron density parameters for the SynGSTC1-GSH complex were transferred from the ELMAM2 database2, which provides parameters averaged over experimental peptide electron densities from ultra-high resolution X-ray scattering data [28]. In addition, polarization effects due to the environment were estimated in the transferred electron density using the procedure described recently by Leduc et al. [29] and implemented in *MoProViewer* software (version 0.1.1302) [30]. The electrostatic interaction energy

( $E_{tot}^{elec}$ ) between the glutathione ligand and the SynGSTC1 active site residues was computed using *Charger*, which is a fast and analytical electrostatic energy calculation tool [31] also implemented in *MoProViewer*.  $E_{tot}^{elec}$  includes two terms, the electrostatic interaction permanent energy  $E_{perm}^{elec}$  and the polarization contribution  $E_{pol}^{elec}$  (hence  $E_{tot}^{elec} = E_{perm}^{elec} + E_{pol}^{elec}$ ). The *MoProViewer* database transfer tool enables an automatic parameter transfer on the structure with appropriate formal charge assignment (+1e for arginine and lysine, −1e for aspartate and glutamate, 0 for others). The His38 and His61 of SynGSTC1 were protonated on the Nε atom and the formal charge of glutathione was set to −1e. The procedure is detailed in the Supplementary Materials.

#### 2.4. Molecular Dynamics Simulation

The Molecular Dynamics simulations presented in this study were based on the crystallographic structure of SynGSTC1 in complex with GSH. This system was immersed in a cubic simulation cell of length equal to 73 Å filled by a solvent of 9881 water molecules with 150 mM NaCl. The simulations were performed using *NAMD 3.0* [32] with the CHARMM36 [33] force field for proteins and the TIP3P water model [34]. The parameters for the GSH ligand were generated by the CHARMM general force field (CGenFF) [35]. Long-range electrostatic forces were evaluated using the particle mesh Ewald algorithm with a grid spacing of 1.0 Å. A smoothed 12.0 Å spherical cutoff was applied to truncate the short-range van der Waals and electrostatic interactions. The temperature was maintained at 300 K thanks to the Langevin thermostat and the pressure at 1 atm thanks to the Langevin piston method. Covalent bonds involving hydrogen atoms were restrained to their equilibrium length by the Rattle algorithm [36] and the water molecules were constrained to their equilibrium geometry using the Settle algorithm [37]. In addition, a mass-repartitioning scheme was used to integrate the equations of motion with a time step of 4 fs, according to Hopkins et al. [38]. A smooth equilibration, along which the positions of the heavy atoms of the protein were restrained harmonically, was carried out during 8 ns before a non-restrained long equilibration of 100 ns. Then, the SynGSTC1-GSH complex was probed in production runs including a long simulation of 500 ns and five independent shorter simulations of 100 ns. These trajectories were visualized and analyzed using *VMD* [39]. These simulations were aimed at exploring the stability of the interactions between the glutathione and the active site as well as the flexibility of the protein interdomain linker.

#### 2.5. Enzymatic Assays

The GSH-conjugation activity was assayed at 25 °C toward 1-chloro-2,4-dinitrobenzene (CDNB), benzyl-isothiocyanate (BITC), 2-phenethyl-isothiocyanate (PITC) or p-nitrophenyl butyrate (PNP-butyrate). The reactions were performed in 500 µL of 30 mM Tris-HCl (pH 8.0) and 1 mM EDTA for CDNB and PNP-butyrate and 100 mM phosphate buffer (pH 6.5) for ITC derivatives in the presence of various concentrations of CDNB (0–4000 µM), BITC (0–1000 µM), PITC (0–1000 µM) or PNP-butyrate (0–2000 µM) at a fixed saturating GSH concentration. Peroxidase and thiol-transferase activities were assayed at 25 °C toward cumene hydroperoxide (CuOOH) and 2-hydroxyethyl disulfide (HED) in a NADPH-coupled spectrophotometric method by following the absorbance at 340 nm. The reactions were carried out in 500 µL of 30 mM Tris-HCl (pH 8.0) containing 200 µM NADPH, 0.5 unit of yeast glutathione reductase and various concentrations of HED (0–500 µM) or CuOOH (0–3000 µM) at a fixed GSH concentration. The optimum pH of the wild-type enzyme and its variants was determined against PITC using 100 mM sodium citrate, phosphate, or borate buffers at pH ranging from 4.0 to 11.0. GSH-conjugation activity was determined as described above.

For all activity assays, the recombinant protein, used at a concentration (3 µM) within the linear response range of the enzyme, was added after 2 min of preincubation and the variation of absorbance monitored using a Cary 50 spectrophotometer. The activity recorded without enzymes was subtracted and three independent experiments were performed at each substrate concentration. The kinetic parameters, apparent  $K_m$  (Michaelis

constant) and  $k_{software\ cat}$  (turnover number) were determined by fitting the data to the nonlinear regression Michaelis–Menten model in *GraphPad Prism* (version 8, GraphPad Software, Inc., San Diego, CA, USA). The  $k_{cat}$  values were expressed as  $\mu\text{mol}$  of substrate oxidized per second per  $\mu\text{mol}$  of enzyme (i.e., the turnover number in  $\text{s}^{-1}$ ) using specific molar absorption coefficients of  $9600\ \text{M}^{-1}\cdot\text{cm}^{-1}$  at 340 nm for CDNB,  $9250\ \text{M}^{-1}\cdot\text{cm}^{-1}$  at 274 nm for BITC,  $8890\ \text{M}^{-1}\cdot\text{cm}^{-1}$  at 274 nm for PITC,  $17700\ \text{M}^{-1}\cdot\text{cm}^{-1}$  at 412 nm for PNP-butyrate and  $6220\ \text{M}^{-1}\cdot\text{cm}^{-1}$  at 340 nm for NADPH.

### 2.6. Phylogenetic Analysis

In total, 222 proteomes of the Cyanobacteria/Melainobacteria group were retrieved from the RefSeq database of the NCBI. These corresponded to 208 proteomes of Cyanobacteria labelled as RefSeq “reference proteomes” or from type strains, and 14 proteomes from noncyanobacterial lineages (i.e., Margulisbacteria, Melainobacteria, Gastranaerophilales) classified in the Cyanobacteria/Melainobacteria group (Table S2). The sequences of the 53 ribosomal protein families (rprots) were retrieved from the 222 proteomes using the riboDB database [40] (Table S3). The corresponding protein sequences were aligned using *MAFFT v7.453* with the accurate option L-INS-I [41]. The resulting multiple alignments were trimmed with *BMGE v1.2* using the BLOSUM30 substitution matrix [42]. The multiple alignments of the 52 rprots present in more than 30% of the 222 analyzed proteomes were combined to build a large supermatrix (222 sequences, 6430 amino acid positions) and used to build a phylogeny using the maximum likelihood method. The tree was inferred with *IQ-TREE* (multicore version 2.2.0 COVID-edition, June 2022) with the LG + C20 + F + R4 evolutionary model [43]. The branch robustness of the inferred tree was computed with the ultrafast bootstrap procedure implemented in *IQ-TREE* (1000 replicates). The resulting tree was rooted using the 14 noncyanobacterial sequences.

The 222 studied proteomes were queried with *BLASTP* using the GSTC1 sequence from the *Synechocystis* sp. PCC 6803 strain (RefSeq protein Id WP\_010873500.1, locus tag SGL\_RS13850) as seed. The 924 GST sequences displaying an E-value lower than  $10^{-3}$  were retrieved and aligned using *MAFFT* with the auto option. A total of 54 partial sequences were discarded from the analysis. A survey of the nr database at the NCBI identified 11 sequences of GSTC in noncyanobacterial bacteria. These 11 sequences were added to the cyanobacterial GSTC sequences. The 881 GSTC sequences were realigned with *MAFFT* with the L-INS-I option and trimmed using *BMGE* with the BLOSUM30 substitution matrix. The 104 kept amino acid positions were used to infer a phylogeny using *FastTree v2* [44] with 20 rate categories of sites, the gamma optimization option, and the Le and Gascuel model [45]. The branch robustness of the inferred tree was estimated using the Shimodaira Hasegawa test (resampling the site likelihoods 1000 times). Finally, a phylogenetic analysis of the 147 cyanobacterial GSTC sequences was performed using *FastTree* and the same parameters (147 sequences, 110 amino acid positions).

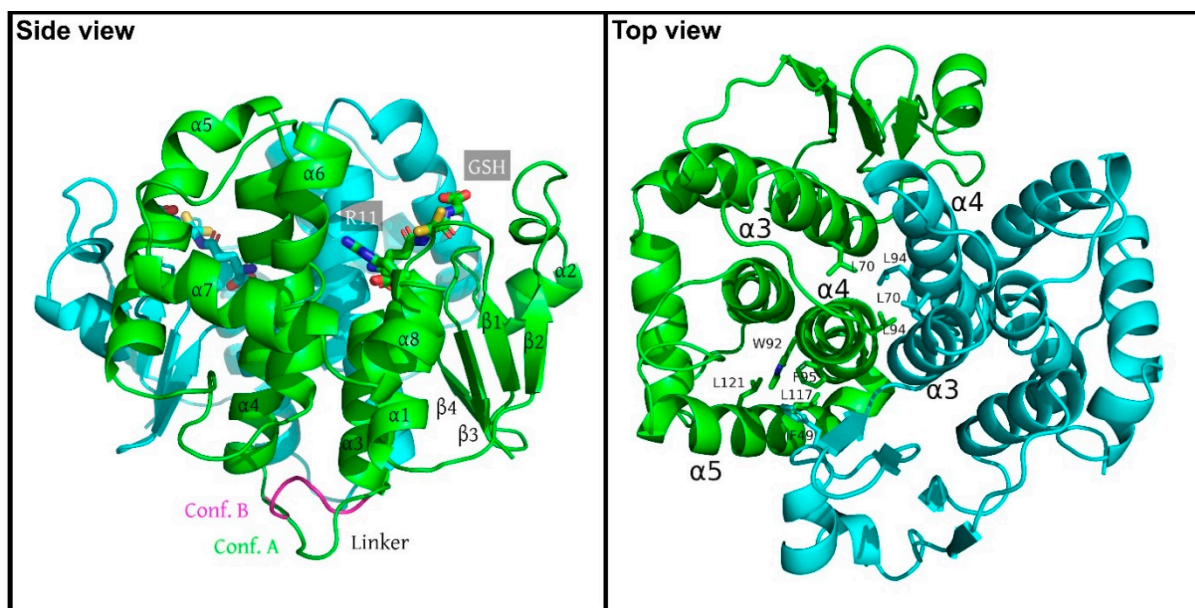
The trees were drawn using *iToL v6.5.8* [46].

## 3. Results and Discussion

### 3.1. Crystal Structure of SynGSTC1

In this study, the crystal structure of the glutathione transferase chi1 from *Synechocystis* sp. PCC 6803 (SynGSTC1) in complex with GSH is presented. We also solved the structures of two variants (S10T and R11A variants in complex with GSH) which did not show significant differences from the wild-type. The protein samples were cocrystallized with an excess of GSH (10:1) in the presence of TCEP to avoid oxidation of the GSH thiol group into sulfenic acid. SynGSTC1 crystallized in space group  $P4_32_12$  with two polypeptide chains in the asymmetric unit. They formed a two-fold dimer that had a globular shape with molecular dimensions of approximately  $55\ \text{\AA} \times 55\ \text{\AA} \times 45\ \text{\AA}$  (Figure 1). The dimer buried  $1710\ \text{\AA}^2$  of surface area for each monomer and was tightly stabilized by ten hydrogen bonds and six salt bridges (Table S4). At the core, a four-helix bundle consisting of the  $\alpha 3$  and  $\alpha 4$  helices of the two monomers buried aliphatic residues (L70 and L94 of chains A and B).

This interaction pattern was complemented by a lock-and-key motif where the F49 residue fitted into a low-polar cavity of the adjacent subunit (W92, F95, L117, L121) (Figure 1).

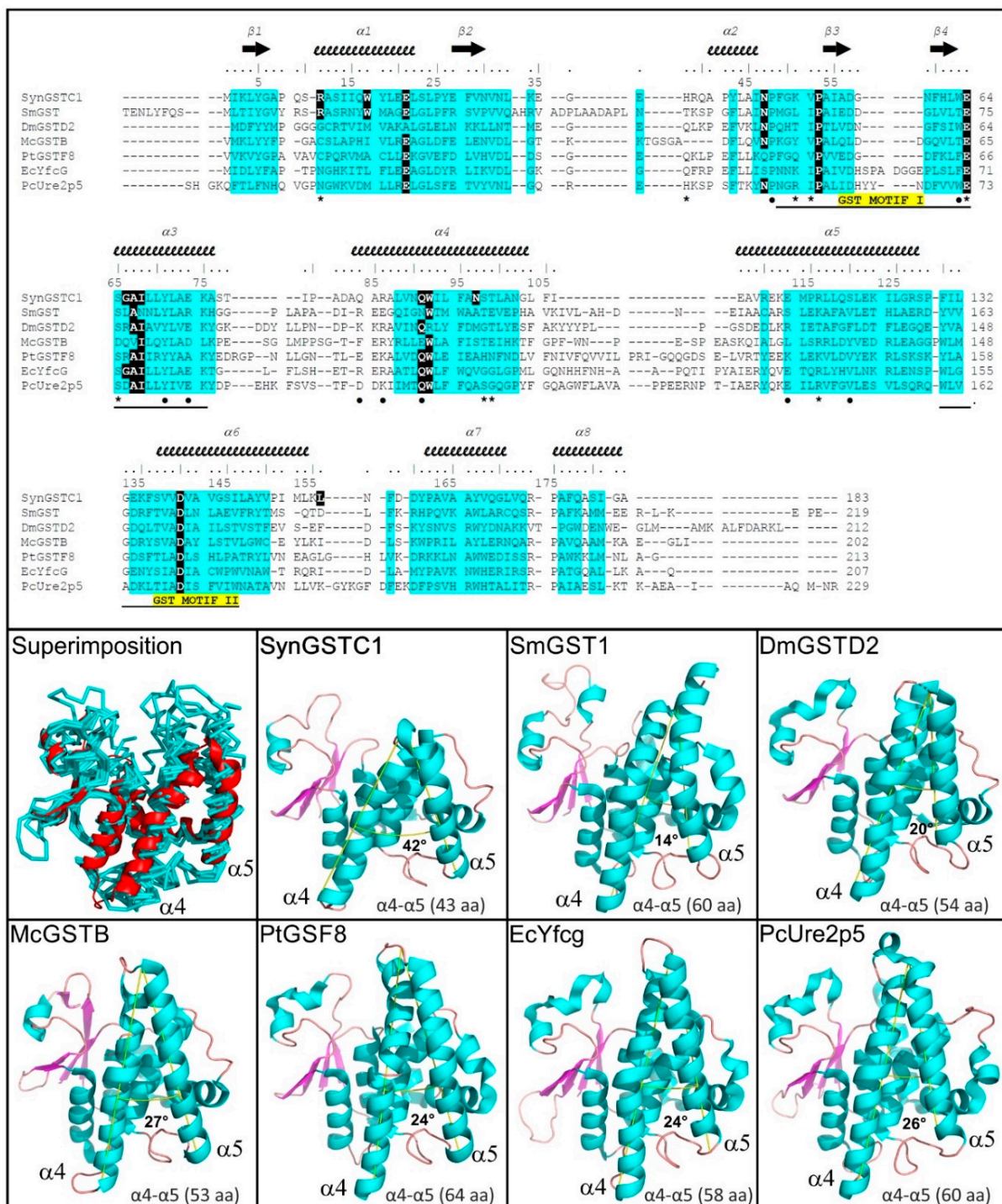


**Figure 1.** Crystal structure of the SynGSTC1 dimer, left, and rotated 90°, right. The monomers A and B are shown in ribbon mode and colored green and blue, respectively. (Left), side view. The secondary structures of the monomer A are labelled. In each monomer, the side chain of residue R11 and the glutathione molecules are labelled and highlighted as sticks. Both conformations of the linker are shown in monomer A. (Right), top view. The figure highlights the hydrophobic patches on SynGSTC1 dimer interface. L70 and L94 of both monomers are buried in the center of the dimer. This interaction pattern is complemented by a lock-and-key motif where the F49 residue (blue) fits into a low-polar cavity of the adjacent subunit (W92, F95, L117, L121) (green). The symmetry related lock-and-key motif is not shown for clarity.

Both subunits were very similar structures and could be superimposed within 0.33 Å root-mean-square deviation over 181  $\alpha$ -carbon atoms. The SynGSTC1 protomer adopted the conserved GST fold that was subdivided into two domains for clarity (N-terminal domain  $\beta 1\alpha 1\beta 2\alpha 2\beta 3\beta 4\alpha 3$  and C-terminal domain  $\alpha 4\alpha 5\alpha 6\alpha 7\alpha 8$ ). As mentioned in the introduction, the chain length of GSTCs (approximately 180 residues) was significantly shorter by at least 20 residues compared to most canonical GSTs [47]. The  $\alpha 4$ – $\alpha 5$  hairpin pattern was significantly shortened (roughly 10 residues) and the angle between these two helices ( $\sim 42^\circ$ ) was twice that usually observed (Figure 2). This “missing” region made the active site of SynGSTC1 very open, with no clear pocket for the hydrophobic substrate (H-site). Both motifs I (47–71) and II (129–147) played their expected structural roles. In motif I, the V52–P53 peptide bond was cis, and V52 formed the typical antiparallel  $\beta$ -sheet-like interaction with the cysteine moiety of GSH. Motif II contained the Ncap sequence  $^{137}\text{SVVD}^{140}$  where the side chains of the serine and aspartic acid residues contributed to the stabilization of the  $\alpha 6$  helix [19]. The linker ( $^{76}\text{ASTIPAD}^{82}$ ) between the N- and C-terminal domains was peculiar because it had no aliphatic or aromatic residue wedged between these two domains as usually observed [48,49]. The consequence was an interdomain linker without a unique conformation. The quality of the electron density allowed the building of two major conformations (Figures 1 and S1). To investigate this property, we performed molecular dynamics simulations of the SynGSTC1-GSH complex in an aqueous environment. The simulation revealed a protein very stable with the linker as one of the most mobile regions. The time-evolution of the  $\varphi$  and  $\psi$  torsion angles of the linker residues revealed transitions between two main conformations during the trajectory



(Figure S2). Interestingly, these two conformations corresponded to those observed in the crystal structure.



**Figure 2.** Comparison of SynGSTC1 with structural homologs. The top figure shows a structure-based sequence alignment, and the bottom figures highlight that SynGSTC1 has the shortest  $\alpha 4$ - $\alpha 5$  hairpin and the highest angle between  $\alpha 4$  and  $\alpha 5$  helices. Crystal structure and sequences can be found at the Protein Data Bank (<http://www.rcsb.org>, accessed on 30 September 2022): SynGSTC1, this study, PDB ENTRY 8A18; SmGST, GST from *Sinorhizobium meliloti* 2011, PDB entry 4NHW; DmGSTD2, GST delta 2

from *Drosophila melanogaster*, PDB entry 5F0G; McGSTB, GST beta from *Methylococcus capsulatus* str. Bath, PDB entry 3UAP; PtGSTF8, GST phi 8 from *Populus trichocarpa*, PDB entry 5F07; EcYfcG, GST nu from *Escherichia coli* K-12, PDB entry 5HFK; PcUre2p5, Ure2p 5 from *Phanerodontia chrysosporium*, PDB entry 4F0C. The characteristics of the top figure are as follows: secondary structures are labelled and shown using arrows ( $\beta$ -strands) and squiggles (helices); common regions, i.e., regions with no gaps and with pairwise residue distances less than 4 Å are highlighted in blue; the invariant residues in the GST chi class are in bold type, coloured white and highlighted in black; residues that participates in dimer stabilization of SynGSTC1 via strong polar interactions are marked with •; residues involved in binding glutathione (G-site) in SynGSTC1 are marked with \*. The characteristics of the bottom figures are as follows: the models are shown in the cartoon or ribbon modes; the  $\alpha 4$  and  $\alpha 5$  helices are labelled; the first figure shows a superimposition of the seven structures where SynGSTC1 is colored red and the others cyan; in the other figures, the estimated angle between  $\alpha 4$  and  $\alpha 5$  helices is provided as well as the number of amino acids in the  $\alpha 4$ – $\alpha 5$  hairpin; the angles were calculated using the *AngleBetweenHelices* script (<https://pymolwiki.org/index.php/AngleBetweenHelices>, accessed on 30 September 2022) implemented in *PyMol Molecular Graphics System* (Version 2.0 Schrödinger, LLC, New York, NY, USA).

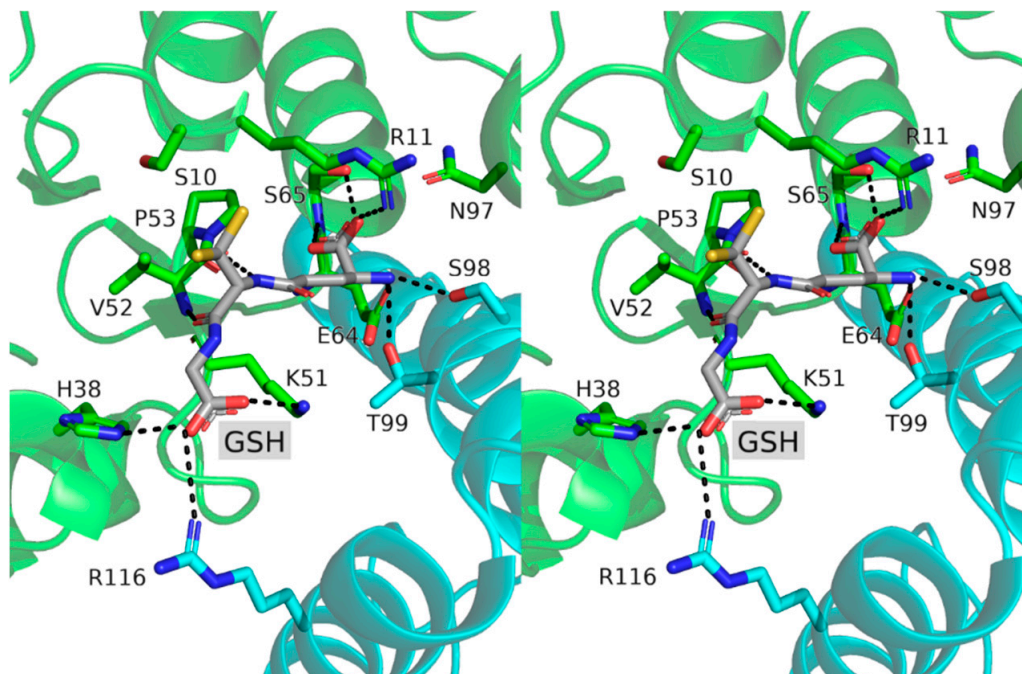
### 3.2. Structural Comparison

A search for the structural homologs using the Dali server (<http://ekhidna2.biocenter.helsinki.fi/dali/>, accessed on 30 September 2022) ranked proteobacterial nu GSTs and fungal GSTs from the Ure2p class at the top of the list [50]. The other hits included proteobacterial beta GSTs, insect delta GSTs, an unclassified proteobacterial GST and plant phi GSTs. To better depict the proximities of these structures, an additional multiple structural alignment was performed using the mTM-align server (<https://yanglab.nankai.edu.cn/mTM-align/>, accessed on 30 September 2022) [51] (Figure 2). The resulting dendrogram based on the pairwise alignment scores (Figure S3) showed a distribution of the proteins into two clades, one of which containing SynGSTC1 and the unclassified proteobacterial GST (GST SMC00097 from *Sinorhizobium meliloti* 2011, PDB entry 4nhw). The latter had therefore the most similar structure to SynGSTC1. SMC00097 had one of the structural attributes of SynGSTC1, namely a SRAS motif at the beginning of the  $\alpha 1$  helix in its active site (see below) (Figure 2). The first serine residue adopted the same orientation and did not participate in the stabilization of GSH while the arginine residue did (Figure S4). The closeness between SynGSTC1 and SMC00097 could be explained in a more comprehensive way by a domain-by-domain comparison. Indeed, the overall structures (i.e., both the N-ter and C-ter domains) of SynGSTC1 and SMC00097 overlapped well (Table S5). The proximity of SynGSTC1 with other hits (nu, beta, delta and phi GSTs) was rather due to the good overlap of N-terminal domains.

### 3.3. Active Site Structure and Its Analysis Using Transferred Multipolar Electron-Density

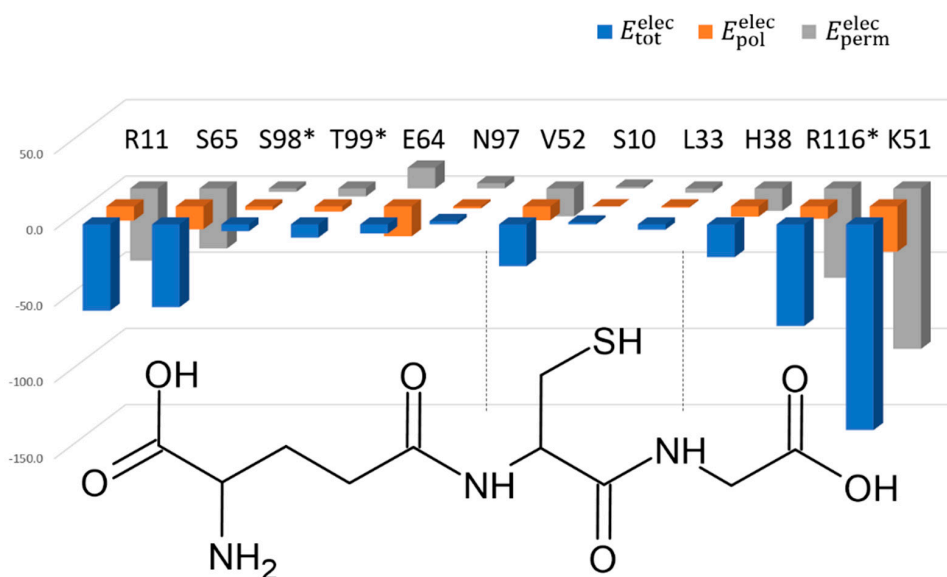
The active site contained GSH tightly bound to the G-site by numerous polar interactions (respectively, six, two and three for the  $\gamma$ -Glu, Cys and Gly moieties) (Figure 3). The GSH Cys moiety adopted two rotamers (**m**,  $\chi_1 = -52^\circ$  and **t**,  $\chi_1 = 172^\circ$ ) exposing the GSH thiol group towards the solvent (Figure 1). The three regular rotamers (**p**, **m**, and **t**) of the glutathione cysteine moiety were accessible during the molecular dynamics simulations with the frequencies of 0.35, 0.42 and 0.16, respectively (Figure S5). The crystal structure did not reveal a strong polar interaction between the sulfur atom of GSH and the enzyme. The smallest distance was 3.8 Å with the amide group of R11. The Y5 residue, recently proposed as a catalytic residue [20], was far too distant to stabilize the GSH-thiolate group during catalysis as the Y5 hydroxyl group was 17 Å away from the GSH sulfur-atom. Based on the sequence analysis of SynGSTC1, we could have thought that S10 played an important role in catalysis. Indeed, this serine residue belongs to the  $^{10}\text{SRAS}^{13}$  motif, which is related to the CXXC active-site motif of thioredoxin [52]. The equivalent serine residue in Ser-GSTs (see introduction) is almost always found hydrogen-bonded to the GSH thiol group while this is not the case in SynGSTC1 [53,54]. Indeed, the OG atom invariably retained the same orientation in all structures (wild-type and variants), and was hydrogen bonded to the main chains of A7 and A12. This interaction network remained stable throughout most of

the molecular dynamics simulations showing that S10 was important for the stabilization of the protein backbone. Whatever its conformation, this serine residue never formed a strong interaction with the GSH thiol group during the simulation (Figure S6).



**Figure 3.** Stereoview of the glutathione binding site of SynGSTC1. The monomers A and B are shown in cartoon mode and colored green and blue, respectively. GSH and residues around it are shown as sticks and labelled. Numbering of residues is according to sequence of SynGSTC1. Strong intermolecular interactions are shown as dashed sticks.

The description of the interactions between a ligand and a protein is most often summarized by the list of residues involved, without quantifying the importance of each. We developed recently a fast and analytical procedure to estimate the electrostatic contribution of each residue to the ligand binding, based on a continuous distribution of electron density of experimental origin [31]. This method implemented in *MoProViewer* [30] was applied on SynGSTC1 in complex with GSH where the contributions of eleven residues were evaluated (distance cutoff of 3.5 Å away from GSH). This included eight residues from one chain (S10, R11, L33, H38, K51, V52, E64, S65 and N97) and three from the other (S98, T99 and R116). *MoproSuite* calculates electrostatic interaction energies  $E_{\text{tot}}^{\text{elec}}$  that are divided into two contributions: a permanent electrostatic interaction term,  $E_{\text{perm}}^{\text{elec}}$ , and a polarization one,  $E_{\text{pol}}^{\text{elec}}$ , which can be interpreted as a molecular recognition term and an adaptation term, respectively (Figure 4, Table S6). By definition, the polarization term is negative and makes the total interaction energy more favorable for all the active site residues and especially for charged residues [29]. We performed the calculations for the two GSH thiol orientations observed in the crystal structure. The orientation of the thiol group did not affect notably the GSH binding, from an electrostatic and dipolar-induction point of view (Table S6). Thus, the following analysis did not depend on the GSH conformation.



**Figure 4.** Permanent, polarization and total electrostatic interaction energies. The permanent  $E_{\text{perm}}^{\text{elec}}$ , polarization  $E_{\text{pol}}^{\text{elec}}$  and total  $E_{\text{tot}}^{\text{elec}}$  electrostatic interaction energies between the glutathione ligand and twelve residues of the SynGSTC1 active site are presented in kcal/mol.  $E_{\text{perm}}^{\text{elec}}$  is computed using the electron density model transferred on the glutathione and the protein atoms, whereas  $E_{\text{tot}}^{\text{elec}}$  is obtained after the electron density polarization procedure. Finally,  $E_{\text{pol}}^{\text{elec}}$  is computed using  $E_{\text{pol}}^{\text{elec}} = E_{\text{tot}}^{\text{elec}} - E_{\text{perm}}^{\text{elec}}$ , and represents the polarization contribution to the total electrostatic interaction energy. The reported energy values have been averaged over the two conformations of the glutathione (A and B) observed in the crystal structure and over the two monomers. The GSH formula has been added to highlight the proximity of the residues to the GSH moieties. The residues marked with a star (\*) in the figure are not from the same monomer as glutathione. The numerical values of these energies and the associated standard deviations are available in the Supplementary Materials.

The permanent interaction energy  $E_{\text{perm}}^{\text{elec}}$  pictures the electrostatic complementarity between the GSH chemical groups and the residues lining the binding site. GSH was assumed to bear three charges: a zwitterionic  $\gamma$ -glutamic acid moiety and a terminal glycine carboxylate group. The SynGSTC1 residues with the largest contributions were R11, K51, R116, which formed salt bridges with the GSH negative charges (Figure 3, Table S6). As an example, the energy values  $E_{\text{tot}}^{\text{elec}}$ ,  $E_{\text{perm}}^{\text{elec}}$  and  $E_{\text{pol}}^{\text{elec}}$  obtained for R11 were  $-56.6 \text{ kcal}\cdot\text{mol}^{-1}$ ,  $-47.3 \text{ kcal}\cdot\text{mol}^{-1}$  and  $-9.3 \text{ kcal}\cdot\text{mol}^{-1}$ , respectively. S65 showed the most favorable  $E_{\text{perm}}^{\text{elec}}$  among the uncharged residues ( $E_{\text{perm}}^{\text{elec}} = -39.3 \text{ kcal}\cdot\text{mol}^{-1}$ ), and its contribution was close to those of R11 and R116 when the dipolar induction is included ( $E_{\text{tot}}^{\text{elec}} = -54.3 \text{ kcal}\cdot\text{mol}^{-1}$ ). This serine residue was double-hydrogen-bonded to the  $\gamma$ -Glu carboxylate group. This interaction pattern, well conserved in GSTs, is ensured either by a serine residue or a threonine residue [55]. The negatively charged E64 residue was an interesting case because it had an unfavorable  $E_{\text{perm}}^{\text{elec}}$  ( $13.6 \text{ kcal}\cdot\text{mol}^{-1}$ ) that underwent a significant dipolar induction ( $E_{\text{pol}}^{\text{elec}} = -19.7 \text{ kcal}\cdot\text{mol}^{-1}$ ) to interact with the positively charged N-terminal amine group of GSH ( $E_{\text{tot}}^{\text{elec}} = -6.0 \text{ kcal}\cdot\text{mol}^{-1}$ , Figure 4, Table S6). The major contributors for the GSH  $\gamma$ -Glu moiety,  $E_{\text{perm}}^{\text{elec}}$  speaking, were therefore R11 via its guanidium group and S65 via its amide and hydroxyl groups (Figure 3). This showed that the site where the zwitterionic fragment of GSH was located was an electrophilic site. This property is verified in the crystallographic structures of glutathione-free GSTs because they often contain a negative ion in this site such as chloride, acetate or formate ions [56]. In addition, this electrophilic site was found to be catalytically important as it hosts the  $\gamma$ -Glu carboxylate group which is presumed to decrease the pKa of the GSH thiol group [57]. The glycine part of GSH was surrounded by the two positively charged

K51 and R116 residues, and by the lateral chain of H38 residue. These residues tightly stabilized the GSH C-terminal carboxylate group (Figures 3 and 4). Finally, the GSH Cys part was strongly stabilized by a single residue (V52) via two main-chain–main-chain hydrogen bonds (Figure 3). This twofold contribution was significantly lower compared to that of S65 probably because the V52-GSH interaction did not involve charged groups. The S10 residue, which was assumed to be the catalytic residue interacting with the thiol group, presented an unfavorable electrostatic interaction energy and did not contribute to the GSH stabilization ( $E_{\text{tot}}^{\text{elec}} = 1.3 \text{ kcal}\cdot\text{mol}^{-1}$ , Table S6). It also showed an almost zero polarization energy so this residue was not affected by the binding of the glutathione. This correlated well with the fact that the crystal structure of SynGSTC1 revealed no intermolecular interaction between S10 and GSH. The side chain of the “main” tyrosine residue of TyrGSTs (Tyrosine type GSTs) was always observed interacting with the thiol group of GSH in the crystal structures. The “main” serine residue of SerGSTs plays the same role in most known structures. We evaluated the electrostatic contribution of residues to GSH binding in a TyrGST (and a SerGST) containing a putative hydrogen bond between the tyrosine (serine) residue and GSH (Table S6). In both cases, the important residue (tyrosine or serine) provided a stabilizing effect on the ligand ( $E_{\text{tot}}^{\text{elec}} = -12 \text{ kcal}\cdot\text{mol}^{-1}$  and  $E_{\text{tot}}^{\text{elec}} = -7.3 \text{ kcal}\cdot\text{mol}^{-1}$ , respectively, Table S6). However, this contribution was never predominant. The main anchor points remained the positively charged residues that stabilized the terminal carboxylate groups of GSH.

### 3.4. Biochemical Characterization of SynGSTC1 and Variants

We recently detected an activity for SynGSTC1 toward methylglyoxal as substrate and also tested glutathione transferase reactions namely aromatic substitution, and addition using, respectively, 1-chloro-2,4-dinitrobenzene (CDNB) and isothiocyanates (ITCs) as substrates [15]. In addition to these activities, we tested here the ability of SynGSTC1 to conjugate GSH on 4-nitrophenyl butyrate (PNP-butyrates) by transacylation and to reduce hydroperoxide toward cumene hydroperoxide (CuOOH). The measured activities ( $k_{\text{cat}}/K_{\text{m}}$ ), respectively of  $49.0 \pm 1.7 \text{ M}^{-1}\cdot\text{s}^{-1}$  for PNP-butyrates and  $604.6 \pm 37.7 \text{ M}^{-1}\cdot\text{s}^{-1}$  for CuOOH are similar to the one measured toward CDNB  $112.5 \pm 14.2 \text{ M}^{-1}\cdot\text{s}^{-1}$ . These activities remain significantly lower than those measured with ITCs ( $6.7 \times 10^5 \pm 0.2 \times 10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$  and  $5.7 \times 10^5 \pm 0.2 \times 10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$  for PITC and BITC, respectively) due to a higher affinity of the enzyme for PITC and BITC ( $31.4 \pm 3.5$  and  $82.0 \pm 10.0 \mu\text{M}$ , respectively) associated to a higher turn-over number ( $21.0 \pm 0.5 \text{ s}^{-1}$  and  $45.0 \pm 1.6 \text{ s}^{-1}$ , respectively) (Table S7).

We also investigated the structure–activity relationships of SynGSTC1 by targeting the first two residues of the  $^{10}\text{SRAS}^{13}$  motif, S10 being suspected to activate glutathione as in Ser-GSTs and R11 because of its ubiquity in GSTCs (see below in Section 3.5). The kinetic constants and the effect of pH on activities toward PITC were determined for S10T, S10A, S10C and R11A variants (Table 2 and Figure S7). The optimal pH of SynGSTC1 WT (7.4 units) was in the same range as usually observed for GSTs [57]. The substitution of S10 by a threonine residue slightly decreased the optimal pH of the enzyme (6.9 vs. 7.4 for WT) and the catalytic efficiency ( $k_{\text{cat}}/K_{\text{m}}$ ) of the protein ( $2.1 \times 10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$  for S10T vs.  $3.7 \times 10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$  for WT). This result was consistent with the crystal structure of S10T which was superimposable to the wild-type (Figure S8). The bulkier threonine side chain did not impair GSH binding. Indeed, the GSH apparent affinity ( $K_{\text{m}}$ ) was not altered in the S10T variant (Table 2). Furthermore, a sequence analysis of GSTCs (see below in Section 3.5) showed either a serine or a threonine as the first residue of the active site motif ( $^{10}\text{SRAS}^{13}$  in SynGSTC1). S10A remained effective even though the decrease was greater than for S10T, being divided by 10 and 1.2 in S10A and S10T, respectively, compared to WT. All the kinetic parameters were affected roughly similarly. The crystal structure of SynGSTC1 did not show interaction between S10 and GSH. Instead, S10 was rather involved in stabilizing the  $\beta 1$ - $\alpha 1$  loop in the close vicinity of the G-site (see above in Section 3.3) suggesting that the S10A substitution most likely disrupted the integrity of the active site. This resulted in a degradation of the catalytic constants and a moderate increase of the

catalysis optimal pH (shift of 0.4 unit compared to WT). The R11A substitution also did not fully abolish the activity of the enzyme even though it decreased significantly (divided by a factor close to 250 as compared to WT). R11 formed a salt bridge with the N-terminal carboxylate group of GSH in the crystal structure (see above in Section 3.3). This interaction did not seem to be essential for the catalysis because the GSH apparent affinity in R11A was not much more degraded than in S10A (four and five times higher in S10A and R11A variants, respectively, compared to WT). The electrostatic influence of R11 on the catalytic process was, however, clear since the catalytic rate was 75 times lower in R11A compared to WT. This was accompanied by a significant one-unit increase in optimal pH suggesting a higher GSH-thiol pKa in the R11A variant than in the WT enzyme. These variations appeared small compared to those observed in eta GSTH1-1 from *Agrobacterium tumefaciens*, which harbored an arginine residue at the same position as in SynGSTC1. Indeed, the R34A mutation in AtuGSTH1-1 had a detrimental effect on the catalytic constant, which dropped by at least a factor of 5000 [58]. Finally, the substitution of S10 by a cysteine residue, had the same global effect as the S10A mutation ( $2.9 \times 10^4 \text{ M}^{-1}\cdot\text{s}^{-1}$  for S10C vs.  $3.6 \times 10^4$  for S10A). Unlike the WT protein and other variants, the S10C enzyme was also active ( $k_{\text{cat}}/K_m$  of  $3.36 \times 10^3 \pm 0.08 \times 10^3 \text{ M}^{-1}\cdot\text{s}^{-1}$ ) with HED, a substrate commonly used to characterize Grxs and cysteinyl-GSTs, indicating that this variant acquired a significant thiol-transferase activity.

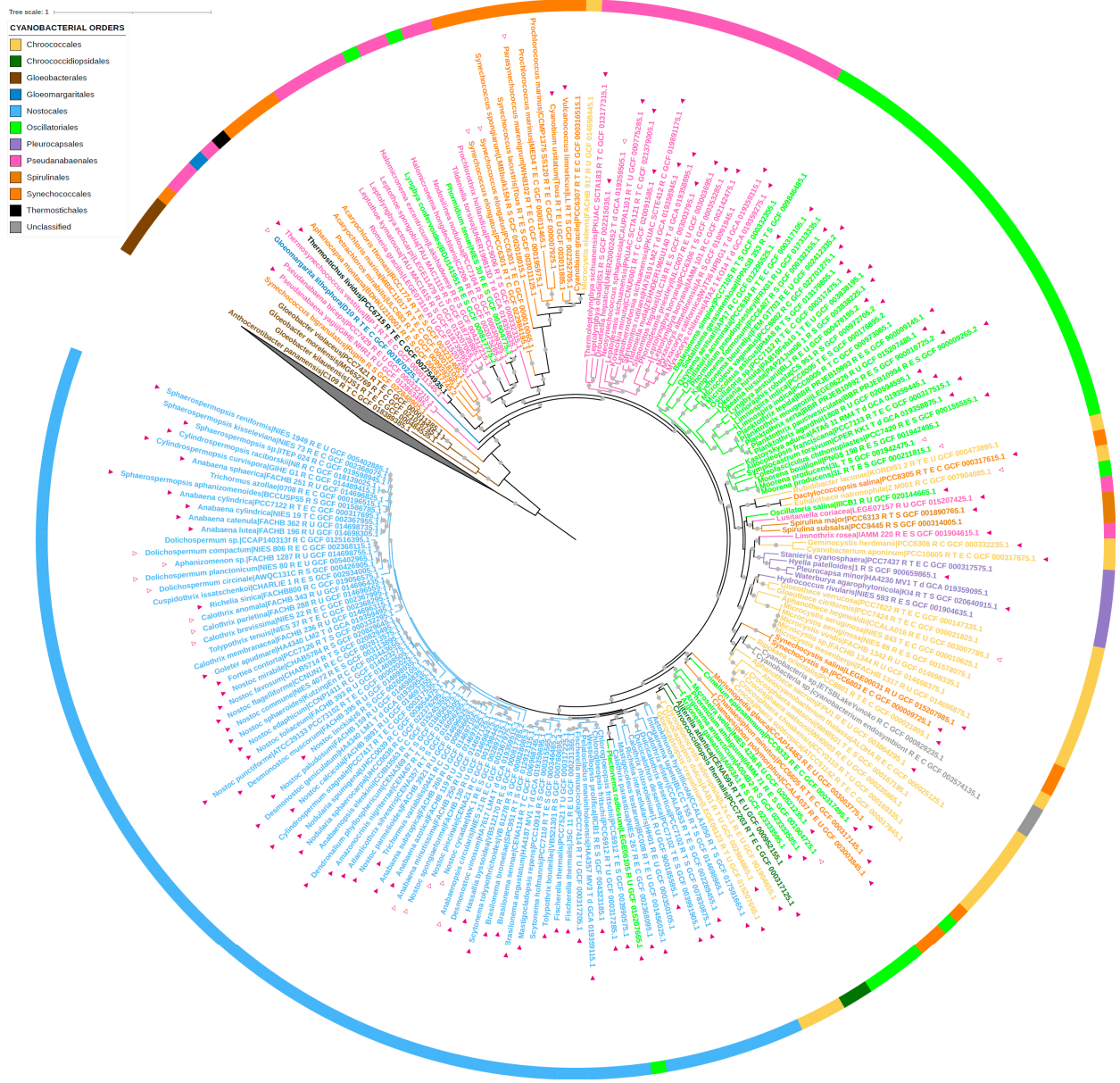
**Table 2.** Kinetic parameters of SynGSTC1 toward model substrates.

	PITC	GSH	HED
<b><math>k_{\text{cat}}</math> (<math>\text{s}^{-1}</math>)</b>			
WT	$12.6 \pm 0.2$		ND
S10T	$7.2 \pm 0.1$		ND
S10A	$2.60 \pm 0.05$		ND
S10C	$1.19 \pm 0.02$		$0.0111 \pm 0.0002$
R11A	$0.170 \pm 0.003$		ND
<b><math>K_m</math> (<math>\mu\text{M}</math>)</b>			
WT	$33.8 \pm 2.5$	$135.2 \pm 7.9$	ND
S10T	$33.6 \pm 2.6$	$142.8 \pm 14.6$	ND
S10A	$89.7 \pm 6.4$	$528.4 \pm 33.0$	ND
S10C	$33.0 \pm 3.0$	$2149 \pm 123$	$3.3 \pm 0.4$
R11A	$108.4 \pm 6.0$	$719.2 \pm 58.2$	ND
<b><math>k_{\text{cat}}/K_m</math> (<math>\text{M}^{-1}\cdot\text{s}^{-1}</math>)</b>			
WT	$3.73 \times 10^5 \pm 0.06 \times 10^5$		ND
S10T	$2.14 \times 10^5 \pm 0.04 \times 10^5$		ND
S10A	$2.89 \times 10^4 \pm 0.06 \times 10^4$		ND
S10C	$3.60 \times 10^4 \pm 0.06 \times 10^4$		$3.36 \times 10^3 \pm 0.08 \times 10^3$
R11A	$1.53 \times 10^3 \pm 0.02 \times 10^3$		ND

The apparent  $K_m$  values of SynGSTC1 wild-type and variants (S10T, S10A, S10C and R11A) were determined by varying substrate concentrations at a fixed saturating GSH concentration. The apparent  $K_m$  and  $k_{\text{cat}}$  values were calculated with *Prism 8* software using the Michaelis–Menten equation as nonlinear regression model. Results are means  $\pm$  S.D. ( $n = 3$ ).

### 3.5. Comparative Genomic Analysis

A similarity-based survey of 222 reference proteomes of the Cyanobacteria/Melainobacteria group led to the identification of 870 full-length GSTC1 homologues (BLASTP E-value cutoff  $10^{-3}$ ). The phylogenetic analysis of these sequences led to a large tree (Figure S9). According to this tree, the glutathione transferase chi1 from *Synechocystis* sp. PCC 6803 (SynGSTC1) belonged to a large group of 147 sequences displaying a SRAS motif or related motifs (Figures S9 and S10 and Table S8). These 147 GSTC protein sequences displayed more than 35% of sequence identity and were largely distributed in Cyanobacteria, being present in 144 of the 208 analyzed cyanobacterial proteomes (Figures 5 and S11 for high-quality version). In contrast, they were absent in the noncyanobacterial members of the Cyanobacteria/Melainobacteria group. More precisely, they were present in all cyanobacterial orders excepted Gloeobacterales, the oldest branching extant group of cyanobacteria [59].



**Figure 5.** Phylogeny of the 222 proteomes of Cyanobacteria/Melainabacteria group considered in this study. The tree was inferred with *IQ-TREE* using the 52 rprots sequences present in more than 70% of the 222 proteomes (6430 amino acid sites, LG + C20 + F + R4 evolutionary model). The scale bar corresponds to the average number of substitutions per site. Gray circles correspond to ultrafast bootstrap values >90% (1000 replicates). The taxonomy of each proteome is indicated: Gloeobacteriales (brown), Synechococcales (orange), Pseudanabaenales (pink), Gloemargaritales (dark blue), Thermostichales (black), Oscillatoriales (light green), Chroococcales (yellow), Pleurocapsales (purple), Chroococciopsidales (dark green), Nostocales (light blue), and unclassified (gray). The 122 GSTC protein sequences harboring the SRAS motif are indicated with filled triangles, while the 25 GSTC sequences harboring variants of the SRAS motif are indicated with empty triangles. All the motifs are described in the Supplementary Table S8. The phylogeny of these 147 GSTC sequences is shown as Figure S10. A high-quality pdf version of the tree is provided as Figure S11 in the online Supplementary Materials.

To go further, we inferred the phylogeny of the 147 sequences displaying the SRAS motif (or related motifs) (Table S9). As expected, due to the restricted number of amino acid positions retained after the alignment trimming, branch supports were overall low

(Figure S10). Despite this global lack of support, the resulting tree showed clearly that sequences harboring the SRAS motif and sequences harboring related motifs were mixed on the tree, indicating that the canonical SRAS motif was lost several times independently during the diversification of GSTCs. Furthermore, the topology of the tree also showed some inconsistencies with the phylogeny of species (Figures 5 and S10). For instance, some Chroococcales sequences emerged within Nostocales (Figure S10), indicating that the evolutionary history of GSTCs harboring the SRAS (or related motifs) was impacted by horizontal gene transfers (HGTs) (Figure S10). Interestingly, these HGTs also contributed to spread GSTC1 outside of Cyanobacteria, since homologues were found in a few non-cyanobacterial bacteria (Figure S9). Most of them were marine and freshwater bacteria and some were recently closely related to cyanobacteria as Planctomycetaceae bacterium TMED241. Indeed, it was found that this bacterium contains a circadian clock *kaiABC* operon, which is typically found in cyanobacteria [60].

The majority of the 147 GSTC sequences had a length of less than 190 amino acids (Table S10). A dozen had longer sequences because they contained extensions at the N-terminus and/or between the secondary structures. All GSTCs had a reduced C-ter domain with a shortening of helices  $\alpha 4$  and  $\alpha 5$  as observed in the crystal structure of SynGSTC1. The SRAS motif ( $^{10}\text{SRAS}^{13}$  in SynGSTC1) was well conserved. The arginine residue was invariant, the first position was replaced in a few cases by a threonine residue and the last two positions were a bit more variable. Surprisingly other residues involved in the structural attributes of SynGSTC1 were not conserved such as the patch of leucine residues (L70 and L94) in the core of the dimer, or the key residue of the lock and key motif (F49), or quaternary contributors to the stabilization of GSH (S98, T99, R116) (Figure 2). The sequence alignment revealed the conservation of 13 residues, most of which were located in the N-terminal domain (eight residues) and more precisely in the domain I (six residues) (Tables S8 and S9). The N-ter domain is generally better conserved than the C-ter domain because it contains an extended part of the active site [2]. In one subunit, the set of conserved residues was not centered on the active site but rather on the center of gravity of the monomer. The residues were distributed almost homogeneously around this center and most of them were located at a distance of less than 10 Å from it (Figure S12). This distribution was consistent with what is usually observed in proteins, namely that the most conserved positions tend to be situated in the core of the protein or on functional surfaces [61]. While the structural role of these conserved residues is obvious, it is difficult to identify those that form the signature of GSTs chi and most of them have been shown conserved in a class of GSTs. Only N97 seemed specific to the GST chi class; it most likely contributed electrostatically to the active site, as it was located near the  $\gamma$ -Glu moiety of GSH and close to the guanidinium group of the SRAS motif (Table S9).

#### 4. Conclusions

This study increased the knowledge on the biochemical characteristic acquired on the chi class of GSTs (GSTCs) and detailed for the first time the structural attributes of this GST class, specific to cyanobacteria. These short-sequence GSTs (~180 aa) had a three-dimensional structure with a very open active site because the  $\alpha 4$  and  $\alpha 5$  helices were significantly shorter than those usually observed. The glutathione substrate was tightly bound to the enzyme with its reactive center exposed to the solvent. The transfer of multipolar density parameters from small peptides to SynGSTC1 permitted the gradation of residues involved in GSH stabilization. The two carboxylate groups of GSH were the two chemical groups that best adhered to the protein.

GSTCs contained a SRAS conserved motif at the N-terminus of the  $\alpha 1$  helix indicating that they belonged to the SerGST group because the first residue of the motif was a serine residue. However, this serine residue was not directly involved in the catalytic act as assumed in SerGSTs [9]. The SRAS motif appeared to constrain the conformation of the serine side chain towards the interior of the protein and not towards the thiol group of GSH. S10 (in SynGSTC1) had a weak and unfavorable electrostatic influence on GSH and its



mutation did not drastically alter the catalytic properties of the enzyme. The denomination TyrGST, CysGST, SerGST and AtyGST (tyrosine type GST, . . . , Atypical GST) has the advantage of simplifying the confusing and cumbersome Greek letter classification. It is relevant in the case of TyrGSTs from a phylogenetic point of view [1]. It is also appropriate in the case of CysGSTs because the cysteine residue is covalently bound to the substrate in one step of the catalytic mechanism [62]. The disadvantage of the residue-based naming is its stigmatization on one residue that may not have a strong link to the activity of the enzyme as is the case for SynGSTC1.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biom12101466/s1>; Table S1. List of the PCR and mutagenic primers used in this study; Table S2. List of the 222 studied proteomes; Table S3. List of the 53 ribosomal protein families present in bacteria according to riboDB database; Table S4. Strong intersubunit contacts in SynGSTC1; Table S5. Comparison of SynGSTC1 with its structural homologs; Table S6. Permanent, polarization and total electrostatic interaction energies in the active sites of SynGSTC1, of Epsilon 2 GST from *Anopheles Gambiae* (AgGSTE2) and of Alpha 1 GST from chicken (GgGSTA1); Table S7. Kinetic parameters of SynGSTC1 toward model substrates; Table S8. List of the 147 full-length protein sequences displaying a SRAS motif (or a related motif) identified in the 222 proteomes of Cyanobacteria/Melainabacteria group; Table S9. Invariant amino acid residues in the GST Chi Class; Table S10. Multiple sequence alignment of the 147 GSTCs (displaying a SRAS motif or a related motif) identified in the 222 proteomes of Cyanobacteria/Melainabacteria group; Figure S1. Stereoviews of the 2mFo-DFc map of the SynGSTC1 inter-domain linker; Figure S2.  $\Phi$  and  $\Psi$  torsion angles for the inter-domain linker residues in SynGSTC1 during the simulation; Figure S3. Structure-based phylogenetic tree of SynGSTC1 with structural homologs; Figure S4. Stereoview of the comparison of the SRAS motif in SynGSTC1 and in GST SMc00097 from *Simorhizobium meliloti* 2011; Figure S5. N-C $\alpha$ -C $\beta$ -S $\gamma$  torsion angle of glutathione during molecular dynamics simulation of SynGSTC1; Figure S6. Interatomic distances (Å) between  $\gamma$ -oxygen atom of Ser10 and selected atoms during molecular dynamics simulation of SynGSTC1; Figure S7. Optimal reaction pH of SynGSTC1 and variants S10T, S10A, S10C and R11A; Figure S8. Structural comparison of the active site of SynGSTC1 WT with S10T and R11A variants; Figure S9. Phylogeny of the 870 GST sequences identified in the 222 studied proteomes of the Cyanobacteria / Melainabacteria group and the 11 GSTC-related sequences identified in noncyanobacterial bacteria; Figure S10. Phylogeny of the 147 cyanobacterial GST1 sequences harboring the SRAS motif or related motifs; Figure S11. Phylogeny of the 222 proteomes of Cyanobacteria/Melainabacteria group considered in this study (high quality version of the tree provided in Figure 5); Figure S12. Stereoview of the invariant amino acid residues in the GST Chi Class and WebLogos of aligned GSTCs from cyanobacteria. References [63–68] are cited in Supplementary Materials.

**Author Contributions:** Conceptualization, C.D. and A.H.; formal analysis, E.M., L.M., G.M., B.G., C.B.-A., C.D. and A.H.; funding acquisition, C.D. and A.H.; methodology, E.M., L.M., S.M., B.G. and A.H.; supervision, B.G., F.D., C.D. and A.H.; writing—original draft, E.M., C.D. and A.H.; writing—review and editing, E.M., G.M., B.G., F.C., C.C.-C., C.B.-A., C.D. and A.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by a grant from “Agence Nationale pour la Recherche” as part of the “Investissements d’Avenir” program (ANR-11-LABX-0002-01 and ANR-17-CE20-0008-01). This work was supported by the “Institut Jean Barriol”.

**Institutional Review Board Statement:** Not relevant for this study.

**Informed Consent Statement:** Not relevant for this study.

**Data Availability Statement:** PDB data (8AI8, 8AI9 and 8AIB) are made freely available by the wwPDB (<https://www.wwpdb.org/> (accessed on 30 September 2022)).

**Acknowledgments:** We thank LabEx ARBRE and Région Grand Est for funding LM’s PhD thesis. The authors appreciated the access to the “Plateforme de mesures de diffraction X” of the Université de Lorraine. We acknowledge SOLEIL (Gif Sur Yvette, France) and ESRF (Grenoble, France) for providing synchrotron radiation facilities, and we thank the staffs of PROXIMA-1, PROXIMA-2 and BM07 beamlines for assistance.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

### Abbreviations

GST: glutathione transferase; GSTC, chi-class GST; SynGSTC1, GSTC1 from *Synechocystis* sp. PCC 6803; rmsd, root-mean-square deviation; ESRF, European Synchrotron Radiation Facility; PDB, Protein Data Bank; WT, wild-type.

### References

1. Mashiyama, S.T.; Malabanan, M.M.; Akiva, E.; Bhosle, R.; Branch, M.C.; Hillerich, B.; Jagessar, K.; Kim, J.; Patskovsky, Y.; Seidel, R.D.; et al. Large-scale determination of sequence, structure, and function relationships in cytosolic glutathione transferases across the biosphere. *PLoS Biol.* **2014**, *12*, e1001843. [[CrossRef](#)]
2. Allocati, N.; Federici, L.; Masulli, M.; Di Ilio, C. Glutathione transferases in bacteria. *FEBS J.* **2009**, *276*, 58–75. [[CrossRef](#)] [[PubMed](#)]
3. Ma, X.-X.; Jiang, Y.-L.; He, Y.-X.; Bao, R.; Chen, Y.; Zhou, C.-Z. Structures of yeast glutathione-S-transferase Gtt2 reveal a new catalytic type of GST family. *EMBO Rep.* **2009**, *10*, 1320–1326. [[CrossRef](#)] [[PubMed](#)]
4. Ribeiro, A.J.M.; Tyzack, J.D.; Borkakoti, N.; Holliday, G.L.; Thornton, J.M. A global analysis of function and conservation of catalytic residues in enzymes. *J. Biol. Chem.* **2020**, *295*, 314–324. [[CrossRef](#)]
5. Perrot, T.; Schwartz, M.; Saiag, F.; Salzet, G.; Dumarçay, S.; Favier, F.; Gérardin, P.; Girardet, J.-M.; Sormani, R.; Morel-Rouhler, M.; et al. Fungal Glutathione Transferases as Tools to Explore the Chemical Diversity of Amazonian Wood Extractives. *ACS Sustain. Chem. Eng.* **2018**, *6*, 13078–13085. [[CrossRef](#)]
6. Schwartz, M.; Perrot, T.; Aubert, E.; Dumarçay, S.; Favier, F.; Gerardin, P.; Morel-Rouhler, M.; Mulliert, G.; Saiag, F.; Didierjean, C.; et al. Molecular recognition of wood polyphenols by phase II detoxification enzymes of the white rot *Trametes versicolor*. *Sci. Rep.* **2018**, *8*, 8472. [[CrossRef](#)]
7. Allocati, N.; Masulli, M.; Di Ilio, C.; Federici, L. Glutathione transferases: Substrates, inhibitors and pro-drugs in cancer and neurodegenerative diseases. *Oncogenesis* **2018**, *7*, 8. [[CrossRef](#)] [[PubMed](#)]
8. Nianiou-Obeidat, I.; Madesis, P.; Kissoudis, C.; Voulgari, G.; Chronopoulou, E.; Tsafaris, A.; Labrou, N.E. Plant glutathione transferase-mediated stress tolerance: Functions and biotechnological applications. *Plant Cell Rep.* **2017**, *36*, 791–805. [[CrossRef](#)]
9. Sylvestre-Gonon, E.; Law, S.R.; Schwartz, M.; Robe, K.; Keech, O.; Didierjean, C.; Dubos, C.; Rouhler, N.; Hecker, A. Functional, Structural and Biochemical Features of Plant Serinyl-Glutathione Transferases. *Front. Plant Sci.* **2019**, *10*, 608. [[CrossRef](#)]
10. Wikteliu, E.; Stenberg, G. Novel class of glutathione transferases from cyanobacteria exhibit high catalytic activities towards naturally occurring isothiocyanates. *Biochem. J.* **2007**, *406*, 115–123. [[CrossRef](#)]
11. Shehu, D.; Abdullahi, N.; Alias, Z. Cytosolic Glutathione S-transferase in Bacteria: A Review. *Pol. J. Environ. Stud.* **2019**, *28*, 515–528. [[CrossRef](#)]
12. Meux, E.; Prosper, P.; Ngadin, A.; Didierjean, C.; Morel, M.; Dumarçay, S.; Lamant, T.; Jacquot, J.-P.; Favier, F.; Gelhaye, E. Glutathione Transferases of *Phanerochaete chrysosporium*: S-Glutathionyl-p-hydroquinone Reductase Belongs to a New Structural Class. *J. Biol. Chem.* **2011**, *286*, 9162–9173. [[CrossRef](#)] [[PubMed](#)]
13. Masai, E.; Ichimura, A.; Sato, Y.; Miyauchi, K.; Katayama, Y.; Fukuda, M. Roles of the enantioselective glutathione S-transferases in cleavage of beta-aryl ether. *J. Bacteriol.* **2003**, *185*, 1768–1775. [[CrossRef](#)] [[PubMed](#)]
14. Meux, E.; Prosper, P.; Masai, E.; Mulliert, G.; Dumarçay, S.; Morel, M.; Didierjean, C.; Gelhaye, E.; Favier, F. *Sphingobium* sp. SYK-6 LigG involved in lignin degradation is structurally and biochemically related to the glutathione transferase omega class. *FEBS Lett.* **2012**, *586*, 3944–3950. [[CrossRef](#)] [[PubMed](#)]
15. Kammerscheit, X.; Hecker, A.; Rouhler, N.; Chauvat, F.; Cassier-Chauvat, C. Methylglyoxal Detoxification Revisited: Role of Glutathione Transferase in Model Cyanobacterium *Synechocystis* sp. Strain PCC 6803. *mBio* **2020**, *11*, e00882-20. [[CrossRef](#)]
16. Pandey, T.; Singh, S.K.; Chhetri, G.; Tripathi, T.; Singh, A.K. Characterization of a Highly pH Stable Chi-Class Glutathione S-Transferase from *Synechocystis* PCC 6803. *PLoS ONE* **2015**, *10*, e0126811. [[CrossRef](#)]
17. Feil, S.C.; Tang, J.; Hansen, G.; Gorman, M.A.; Wikteliu, E.; Stenberg, G.; Parker, M.W. Crystallization and preliminary X-ray analysis of glutathione transferases from cyanobacteria. *Acta Crystallogr. Sect. F Crystallogr. Commun.* **2009**, *65*, 475–477. [[CrossRef](#)]
18. Allocati, N.; Casalone, E.; Masulli, M.; Ceccarelli, I.; Carletti, E.; Parker, M.W.; Di Ilio, C. Functional analysis of the evolutionarily conserved proline 53 residue in *Proteus mirabilis* glutathione transferase B1-1. *FEBS Lett.* **1999**, *445*, 347–350. [[CrossRef](#)]
19. Dragani, B.; Stenberg, G.; Melino, S.; Petruzzelli, R.; Mannervik, B.; Aceto, A. The Conserved N-capping Box in the Hydrophobic Core of Glutathione S-Transferase P1-1 Is Essential for Refolding. Identification of A buried and Conserved Hydrogen Bond Important for Protein Stability. *J. Biol. Chem.* **1997**, *272*, 25518–25523. [[CrossRef](#)]
20. ShylajaNaciyar, M.; Karthick, L.; Prakasam, P.A.; Deviram, G.; Uma, L.; Prabakaran, D.; Saha, S.K. Diversity of Glutathione S-Transferases (GSTs) in Cyanobacteria with Reference to Their Structures, Substrate Recognition and Catalytic Functions. *Microorganisms* **2020**, *8*, 712. [[CrossRef](#)]
21. Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125–132. [[CrossRef](#)] [[PubMed](#)]

22. Evans, P.R.; Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1204–1214. [[CrossRef](#)] [[PubMed](#)]
23. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)] [[PubMed](#)]
24. Vagin, A.; Lebedev, A. *MoRDa*, an automatic molecular replacement pipeline. *Acta Crystallogr. Sect. A Found. Adv.* **2015**, *71*, s19. [[CrossRef](#)]
25. Smart, O.S.; Womack, T.O.; Flensburg, C.; Keller, P.; Paciorek, W.; Sharff, A.; Vonrhein, C.; Bricogne, G. Exploiting structure similarity in refinement: Automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 368–380. [[CrossRef](#)]
26. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 486–501. [[CrossRef](#)] [[PubMed](#)]
27. Hansen, N.K.; Coppens, P. Testing aspherical atom refinements on small-molecule data sets. *Acta Crystallogr. Sect. A Found. Adv.* **1978**, *34*, 909–921. [[CrossRef](#)]
28. Domagala, S.; Fournier, B.; Liebschner, D.; Guillot, B.; Jelsch, C. An improved experimental databank of transferable multipolar atom models—ELMAM2. Construction details and applications. *Acta Crystallogr. Sect. A Found. Adv.* **2012**, *68*, 337–351. [[CrossRef](#)]
29. Leduc, T.; Aubert, E.; Espinosa, E.; Jelsch, C.; Iordache, C.; Guillot, B. Polarization of Electron Density Databases of Transferable Multipolar Atoms. *J. Phys. Chem.* **2019**, *123*, 7156–7170. [[CrossRef](#)]
30. Guillot, B.; Enrique, E.; Huder, L.; Jelsch, C. MoProViewer: A tool to study proteins from a charge density science perspective. *Acta Crystallogr. Sect. A Found. Adv.* **2014**, *70*, C279. [[CrossRef](#)]
31. Vuković, V.; Leduc, T.; Jelić-Matošević, Z.; Didierjean, C.; Favier, F.; Guillot, B.; Jelsch, C. A rush to explore protein–ligand electrostatic interaction energy with Charger. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2021**, *77*, 1292–1304. [[CrossRef](#)] [[PubMed](#)]
32. Phillips, J.C.; Hardy, D.J.; Maia, J.D.C.; Stone, J.E.; Ribeiro, J.V.; Bernardi, R.C.; Buch, R.; Fiorin, G.; Hémin, J.; Jiang, W.; et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130. [[CrossRef](#)] [[PubMed](#)]
33. Brooks, B.R.; Brooks, C.L., III; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [[CrossRef](#)]
34. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
35. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690. [[CrossRef](#)] [[PubMed](#)]
36. Andersen, H.C. Rattle: A “velocity” version of the shake algorithm for molecular dynamics calculations. *J. Comput. Phys.* **1983**, *52*, 24–34. [[CrossRef](#)]
37. Miyamoto, S.; Kollman, P.A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* **1992**, *13*, 952–962. [[CrossRef](#)]
38. Hopkins, C.W.; Le Grand, S.; Walker, R.C.; Roitberg, A.E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11*, 1864–1874. [[CrossRef](#)]
39. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38. [[CrossRef](#)]
40. Jauffrit, F.; Penel, S.; Delmotte, S.; Rey, C.; de Vienne, D.M.; Gouy, M.; Charrier, J.P.; Flandrois, J.P.; Brochier-Armanet, C. RiboDB Database: A Comprehensive Resource for Prokaryotic Systematics. *Mol. Biol. Evol.* **2016**, *33*, 2170–2172. [[CrossRef](#)]
41. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
42. 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](#)] [[PubMed](#)]
43. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)] [[PubMed](#)]
44. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490. [[CrossRef](#)]
45. Le, S.Q.; Gascuel, O. An improved general amino acid replacement matrix. *Mol. Biol. Evol.* **2008**, *25*, 1307–1320. [[CrossRef](#)]
46. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)]
47. Hayes, J.D.; Flanagan, J.U.; Jowsey, I.R. Glutathione transferases. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 51–88. [[CrossRef](#)]
48. Polekhina, G.; Board, P.G.; Blackburn, A.C.; Parker, M.W. Crystal structure of maleylacetoacetate isomerase/glutathione transferase zeta reveals the molecular basis for its remarkable catalytic promiscuity. *Biochemistry* **2001**, *40*, 1567–1576. [[CrossRef](#)]
49. Wilce, M.C.J.; Parker, M.W. Structure and function of glutathione S-transferases. *Biochim. Biophys. Acta Prot. Struct. Mol. Enzymol.* **1994**, *1205*, 1–18. [[CrossRef](#)]
50. Holm, L.; Laakso, L.M. Dali server update. *Nucleic Acids Res.* **2016**, *44*, W351–W355. [[CrossRef](#)]
51. Dong, R.; Pan, S.; Peng, Z.; Zhang, Y.; Yang, J. mTM-align: A server for fast protein structure database search and multiple protein structure alignment. *Nucleic Acids Res.* **2018**, *46*, W380–W386. [[CrossRef](#)]

52. Jacquot, J.P.; Gelhaye, E.; Rouhier, N.; Corbier, C.; Didierjean, C.; Aubry, A. Thioredoxins and related proteins in photosynthetic organisms: Molecular basis for thiol dependent regulation. *Biochem. Pharmacol.* **2002**, *64*, 1065–1069. [[CrossRef](#)]
53. Dixon, D.P.; Edwards, R. Glutathione Transferases. *Arabidop. Book* **2010**, *8*, e0131. [[CrossRef](#)]
54. Pegeot, H.; Koh, C.S.; Petre, B.; Mathiot, S.; Duplessis, S.; Hecker, A.; Didierjean, C.; Rouhier, N. The poplar Phi class glutathione transferase: Expression, activity and structure of GSTF1. *Front. Plant. Sci.* **2014**, *5*, 712. [[CrossRef](#)]
55. Sheehan, D.; Meade, G.; Foley, V.M.; Dowd, C.A. Structure, function and evolution of glutathione transferases: Implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* **2001**, *360*, 1–16. [[CrossRef](#)]
56. Sylvestre-Gonon, E.; Morette, L.; Vilorio, M.; Mathiot, S.; Boutilliat, A.; Favier, F.; Rouhier, N.; Didierjean, C.; Hecker, A. Biochemical and structural insights on the poplar tau glutathione transferase GSTU19 and 20 paralogs binding flavonoids. *Front. Mol. Biosci.* **2022**, *9*, 9585866. [[CrossRef](#)] [[PubMed](#)]
57. Dourado, D.F.; Fernandes, P.A.; Mannervik, B.; Ramos, M.J. Glutathione transferase: New model for glutathione activation. *Chemistry* **2008**, *14*, 9591–9598. [[CrossRef](#)]
58. Skopelitou, K.; Dhavala, P.; Papageorgiou, A.C.; Labrou, N.E. A glutathione transferase from *Agrobacterium tumefaciens* reveals a novel class of bacterial GST superfamily. *PLoS ONE* **2012**, *7*, e34263. [[CrossRef](#)]
59. Moreira, D.; Tavera, R.; Benzerara, K.; Skouri-Panet, F.; Couradeau, E.; Gerard, E.; Fonta, C.L.; Novelo, E.; Zivanovic, Y.; Lopez-Garcia, P. Description of *Gloeomargarita lithophora* gen. nov., sp. nov., a thylakoid-bearing, basal-branching cyanobacterium with intracellular carbonates, and proposal for *Gloeomargaritales* ord. nov. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 653–658. [[CrossRef](#)]
60. Dvornyk, V.; Mei, Q. Evolution of *kaiA*, a key circadian gene of cyanobacteria. *Sci. Rep.* **2021**, *11*, 9995. [[CrossRef](#)]
61. Halabi, N.; Rivoire, O.; Leibler, S.; Ranganathan, R. Protein sectors: Evolutionary units of three-dimensional structure. *Cell* **2009**, *138*, 774–786. [[CrossRef](#)] [[PubMed](#)]
62. Board, P.G.; Coggan, M.; Chelvanayagam, G.; Eastal, S.; Jermini, L.S.; Schulte, G.K.; Danley, D.E.; Hoth, L.R.; Griffor, M.C.; Kamath, A.V.; et al. Identification, characterization, and crystal structure of the omega class glutathione transferases. *J. Biol. Chem.* **2000**, *275*, 24798–24806. [[CrossRef](#)]
63. Williams, C.J.; Headd, J.J.; Moriarty, N.W.; Prisant, M.G.; Videau, L.L.; Deis, L.N.; Verma, V.; Keedy, D.A.; Hintze, B.J.; Chen, V.B.; et al. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **2018**, *27*, 293–315. [[CrossRef](#)] [[PubMed](#)]
64. DeLano, W.L. Pymol: An open-source molecular graphics tool. *CCP4 Newsl. Protein Crystallogr.* **2002**, *40*, 82–92.
65. Roret, T.; Thuillier, A.; Favier, F.; Gelhaye, E.; Didierjean, C.; Morel-Rouhier, M. Evolutionary divergence of Ure2pA glutathione transferases in wood degrading fungi. *Fungal Genet. Biol.* **2015**, *83*, 103–112. [[CrossRef](#)] [[PubMed](#)]
66. Aceto, A.; Dragani, B.; Melino, S.; Allocati, N.; Masulli, M.; Ilio, C.D.; Petruzzelli, R. Identification of an N-capping box that affects the  $\alpha$ 6-helix propensity in glutathione S-transferase superfamily proteins: A role for an invariant aspartic residue. *Biochem. J.* **1997**, *322*, 229–234. [[CrossRef](#)]
67. Dirr, H.; Reinemer, P.; Huber, R. X-ray crystal structures of cytosolic glutathione S-transferases: Implications for protein architecture, substrate recognition and catalytic function. *Eur. J. Biochem.* **1994**, *220*, 645–661. [[CrossRef](#)] [[PubMed](#)]
68. Crooks, G.E.; Hon, G.; Chandonia, J.M.; Brenner, S.E. WebLogo: A sequence logo generator. *Genome Res.* **2004**, *14*, 1188–1190. [[CrossRef](#)]