

Article

Purification, Characterization, and Structural Studies of a Sulfatase from *Pedobacter yulinensis*

Caleb R. Schlachter¹, Andrea O'Malley², Linda L. Grimes¹, John J. Tomashek¹, Maksymilian Chruszcz^{2,*}
and L. Andrew Lee^{1,*}

¹ Integrated Micro-Chromatography Systems, 110 Centrum Drive, Irmo, SC 29063, USA; caleb.schlachter@imcstips.com (C.R.S.); linda.grimes@imcstips.com (L.L.G.); john.tomashek@outlook.com (J.J.T.)

² Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA; aomalley@email.sc.edu

* Correspondence: chruszcz@mailbox.sc.edu (M.C.); lee@imcstips.com (L.A.L.); Tel.: +1-803-777-7399 (M.C.); +1-888-560-2073 (L.A.L.)

Abstract: Sulfatases are ubiquitous enzymes that hydrolyze sulfate from sulfated organic substrates such as carbohydrates, steroids, and flavones. These enzymes can be exploited in the field of biotechnology to analyze sulfated metabolites in humans, such as steroids and drugs of abuse. Because genomic data far outstrip biochemical characterization, the analysis of sulfatases from published sequences can lead to the discovery of new and unique activities advantageous for biotechnological applications. We expressed and characterized a putative sulfatase (PyuS) from the bacterium *Pedobacter yulinensis*. PyuS contains the (C/S)XPXR sulfatase motif, where the Cys or Ser is post-translationally converted into a formylglycine residue (FGly). His-tagged PyuS was co-expressed in *Escherichia coli* with a formylglycine-generating enzyme (FGE) from *Mycobacterium tuberculosis* and purified. We obtained several crystal structures of PyuS, and the FGly modification was detected at the active site. The enzyme has sulfatase activity on aromatic sulfated substrates as well as phosphatase activity on some aromatic phosphates; however, PyuS did not have detectable activity on 17 α -estradiol sulfate, cortisol 21-sulfate, or boldenone sulfate.

Keywords: sulfatase; formylglycine; crystallography; 4-methylumbelliferyl sulfate; hydrolase; *Pedobacter yulinensis*; arylsulfatase



Citation: Schlachter, C.R.; O'Malley, A.; Grimes, L.L.; Tomashek, J.J.; Chruszcz, M.; Lee, L.A. Purification, Characterization, and Structural Studies of a Sulfatase from *Pedobacter yulinensis*. *Molecules* **2022**, *27*, 87. <https://doi.org/10.3390/molecules27010087>

Academic Editor: Fengqing Yang

Received: 30 November 2021

Accepted: 21 December 2021

Published: 24 December 2021

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



Copyright: © 2021 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

Sulfatases are enzymes found in all biological domains. These enzymes are involved in diverse cellular functions including cell signaling [1], pathogenesis [2], hormone regulation [3], steroid metabolism [4,5] and glycosphingolipid metabolism [6]. Sulfatase nomenclature stems from the type of substrates they are capable of hydrolyzing, e.g., arylsulfatases (small aromatic molecules, EC 3.1.6.1), steryl-sulfatases (steroids; EC 3.1.6.2), glycosulfatases (glucose sulfates, EC 3.1.6.3), N-acetylgalactosamine-6-sulfatases (chondroitin or keratan sulfate, EC 3.1.6.4) and choline sulfatases (choline, EC 3.1.6.6). Arylsulfatases are of medical interest as these sulfatases may be involved in human diseases such as metachromatic leukodystrophy (MLD), Maroteaux–Lamy syndrome (MPSVI), and X-linked Ichthyosis (XLI), all of which are types of multiple sulfatase deficiency in humans (MSD) [7,8]. In addition, arylsulfatases have industrial applications for analyzing steroid sulfates in human urine by mass spectrometry [4,5]. General substrates for sulfatases are compounds or small molecules containing sulfate esters (ROSO₃[−]) or sulfamates (RN(H)SO₃[−]) [9].

Barbeyron et al. [9] suggested that the nomenclature of sulfatase enzyme families in the UniProt database [10] reflects the enzyme function. Under such classification, there are four sulfatase families (S1–S4). The largest family of sulfatases is S1, containing sulfatases that remove sulfate ester groups by a hydrolytic mechanism with the metal co-factor calcium [9]. In addition, S1 family sulfatases are unique compared to the other sulfatase

families because they contain a modified active site of cysteine or serine that is post-translationally oxidized to a (2S)-2-amino-3-oxopropanoic acid—also referred to as C α -formylglycine (FGly) [9,11,12]. The consensus amino acid sequence directing the post-translational modification is (C/S)XPXR [12–14]. This modification is performed by either an anaerobic sulfatase-maturing enzyme (anSME) or aerobic formylglycine-generating enzyme (FGE) [13,15].

We sought to express and characterize a putative sulfatase from *Pedobacter yulinensis* (Accession No. WP_107215127), a Gram-negative, orange bacterium isolated from sandy soil in the district of Yulin, Shaanxi province, China [16]. The *P. yulinensis* sulfatase, or PyuS, contains the conserved CXPXR sulfatase motif as well as residues for binding the calcium co-factor. PyuS has activity not only on the generic sulfatase substrate 4-methylumbelliferyl sulfate, but also on 5-bromo-4-chloro-3-indolyl phosphate and *para*-nitrophenyl phosphate. PyuS was not able to hydrolyze any of the tested steroid sulfates. Furthermore, we were able to obtain three high-resolution crystal structures of PyuS. The research presented herein describes a previously uncharacterized protein from *P. yulinensis* that has both sulfatase and phosphatase activities, thus providing further insight into the evolutionary and divergent paths that certain enzymes can take over time.

2. Results

2.1. PyuS Contains the Signature CXPXR Sulfatase Motif

The amino acid sequence of PyuS was chosen using the protein Basic Local Alignment Search Tool (BLASTp, [17]) and searching with the sequences of better characterized sulfatases such as those from *Pseudomonas aeruginosa*, *Helix pomatia*, and *Homo sapiens* (Figure 1). Based on the sequence alignment, PyuS contains the conserved CXPXR sulfatase motif. Furthermore, PyuS also contains conserved active site residues that are involved in coordinating the calcium ion which is necessary for sulfate ester hydrolysis (Figure 1). These residues are Asp15, Asp16, Asp276, Asn277, and FGly55 (modified Cys) in PyuS.

2.2. Production and Purification of Recombinant PyuS

The gene for PyuS (Accession No. WP_107215127) was synthesized with a C-terminal His₆-tag (residues GSHHHHHH) in a cloning vector pUC57 from BioBasic (Toronto, Canada) and codon-optimized for *Escherichia coli* expression. The gene was excised with the restriction enzyme SapI and ligated into the ATUM (Newark, CA, USA) expression vector pD451-SR, which has an inducible T7 promoter. No soluble PyuS was obtained with this initial construct (data not shown). Using the SignalP server [18], it was determined that the first 24 amino acids of PyuS have a 98% possibility of being a signal peptide for a secretory pathway involving the Sec translocon and signal peptidase I (SPI) [19]. The fragment coding for this putative signal peptide was removed from the DNA sequence for further PyuS expression and characterization experiments. In addition, to ensure that the cysteine conversion to formylglycine modification occurred to completion, we co-expressed an FGE from *Mycobacterium tuberculosis* [15,20]. The FGE was codon optimized for *E. coli* expression, and both FGE and PyuS were cloned into an in-house polycistronic plasmid expression vector. In this plasmid, both genes are under the control of an isopropyl β -D-1-thiogalactopyranoside (IPTG)-inducible T7 promoter.

PyuS co-expressed with FGE was purified by nickel immobilized metal affinity chromatography (IMAC) and size-exclusion chromatography (SEC) with Tris-based buffers. The yield of recombinant PyuS post-purification was 5 mg per 700 mL of Luria broth (LB) culture. The molecular weight (MW) of recombinant PyuS (without N-terminal signal peptide) is 51.3 kDa (459 residues). Protein concentration was measured by absorbance at 280 nm and using a theoretical molar extinction coefficient at 77,810 M⁻¹ cm⁻¹ [21]. Recombinant PyuS has a calculated isoelectric point (pI) of 8.24. For the amino acid sequence numbering convention, we did not include the numbering for the putative signal peptide, i.e., the 25th residue of native PyuS (glutamine) starts as residue number 2 for our

recombinant protein (Figure 1). Purified recombinant PyuS (with intact C-terminal His-tag) was used for further characterization.

PaeS	1	-----MSKRPNFLVIVADDLGFSIDIGAFG-GEIATPNLDALAIAGLRLTDFHTAS-TCSPTRSMLLTGTDHHiAGIGTMAEALT
HpoS	19	DQSSASAGTRQDAGQPNIVFVLADDFGFHDVGVYHG-SEIHTPTLDALSASGVRLNENYVQP-ICTPTRSQQLMSGRYQIHTGLQHQI INSC
HsaS1	37	-----SGAGASRPPLVFLADDLGWNDVGFHG- SRIRTPHLDALAAGGVLLDNYTQP-LCTPSRSQQLLTGRYQIRTGLQHQI IWPC
HsaS2	22	-----HAASRPNIILVMAADDLGIGDPCYGNKTI RTPNIDRLASGGVKLTQH LAASPLCTPSRAAFMTGRYPVRS GMASWSRTGV
PyuS	2	-----QGSRKPNIIIFIMADDLGWGELGSYGNTFNETPNLDRLSAQGMRF TQAYAAAPVCSPTRASIMTGYPARVGI TDFLPEDE
		: *:::~::~***:* : * . * ** : * * : * : : . . *::*: :~::~* * :
PaeS	78	PELEGKPGYEGHLNERVVALLPELLREAGYQ TLM AGKWHLGLK-----PEQTPHARGFERSFSLLPG-----AANHYGFEPYPD
HpoS	107	QPNALPNDSP----- TLADKLKESGYATHMVGKWHLGFYK-----QEYLPWNRGFDTYFGYLNA-----AEDYFNHN----
HsaS1	118	QPSCVPLDEK----- LLPQLLKEAGYTHMVGKWHLGMYSR-----KECLPTRRGFDTYFGYLLG-----SEDIYSHERCTL
HsaS2	102	FLFTASSGGLP---TDEI TFAKLLKQGYSTALIGKWHLGMSCSKTDFCHHPLHHGFNYFYGI SLTNLRDCKPGEQSVFTTGFKRLVF
PyuS	82	KTDRWLDPTKY-----V TLNEALSASGYHTGIVGKWHLDTDFK---LNKGGPKAHGFNEVIGTESE-----
		: . * ** * : ***** . * :~::~:
PaeS	151	ESTPRIILKGTPALYVEDERYLDTLPEGFYSSDAFGDKLLQYLKERDQSRPFFAYL PFSAPHWPLQAPREIVEKYRGRYDAGPEALRQERL
HpoS	169	--VPWRQV--RYLDLRDNNGPVRNETGQYSAHLFTGKAIDVVQSHNTSKPLFLYLAYQSVHAPLEVPKEYEHKYRNI TDK-----
HsaS1	184	IDALNVTR--CALDFRDGEEVATGYKNMYS TNIFTKRAIALITNHPPEKPLFLYLALQSVHEPLQVPEEYLPYDFIQDK-----
HsaS2	188	LPLQIVGVTLTLTAAALNCLGLLHVPLGVFFSLLFLAALILTLFLGFLHYFRPLNCFMRRNYEIIQQPMSYDNL TQRLTVEAAQFIQRN--
PyuS	140	----YIADGDYFFPYSKIASFDKGTANEYL TDRQCAEANAFITR-NREKPFFLYLSLYSVHTRLEAPVQLVEKYKQKFDQKYGTGKAEQF
		: . . : : : : : : : : * . .
PaeS	241	ARLKEGLVEADVEAHPVLAL TREWEALEDEERAKSARAMEVYAAAMVERMDWNI GRVVDYLRROGELDNTFVLFMSDNG-AEGAL 324...
HpoS	245	-----NRRTFAGMV SALDEGVANLTQALKDKGLWNNTVLI FSTDNNGGQIHAG 291...
HsaS1	262	-----NRHHYAGMVSLMDEAVGNVTAALKSSGLWNNTVLI FSTDNNGGQITLAG 308...
HsaS2	276	-----TETPFLLVLSYLHVHTALFSSKDFAGKSQHG VYGDAVEEMDWSVGQILNLLDELRLANDTLIYFTSDQGAHVEEV 350...
PyuS	225	FG-----ANNVRHESAQRDNPWLAAMLES IDTG VGGIMKTLRETGLAENT IIVFFSDNNGGAGKAG 284...
		. :. : * :. : * :~::~* * :~::~*

Figure 1. Amino acid sequence alignment of sulfatases from *P. aeruginosa* (PaeS), *H. pomatia* (HpoS), *H. sapiens* (HsaS1 and HsaS2), and *P. yulinensis* (PyuS). Residues involved in binding calcium are highlighted in green, and the conserved CXPXR sulfatase motif (Cys to FGly post-translational modification) is highlighted in teal. HpoS residues for metal binding are based on homology as there is no crystal structure available for this protein. N-terminal signal peptides or putative signal peptides were removed from the alignment where applicable. * Indicates completely conserved residues.

2.3. PyuS Crystal Structure

Our structural studies show that, despite PyuS having less than 21% sequence identity with the sulfatase sequences in Figure 1, the tertiary structure among them is highly homologous. We were able to obtain three high-resolution crystal structures of recombinant PyuS, most of which were obtained under conditions containing carboxylic acid(s) and had ligand(s) bound to the protein (PDB codes: 7STT, 7STU, 7STV). Figure 2a displays a 1.60 Å crystal structure of recombinant PyuS with calcium bound in the active site (PDB code: 7STT).

Overall, the size of recombinant PyuS is approximately $60 \times 61 \times 63$ Å (as measured in COOT [22]) and has a calculated solvent-exposed surface area of approximately $17,800$ Å² [23]. The protein crystallized in the P3₂21 space group with one protein chain per asymmetric unit for all structures. The PyuS protein chain has an α/β sandwich fold comprised of 15 α -helices and 13 β -strands that form two anti-parallel β -sheets (Figure 2b). Helix 3 (Figure 2a) contains the CXPXR sulfatase motif and was found in the hydrophobic core of the protein. It is flanked by a β -sheet 1 (β -strands 1–9), which was also found in the protein core, and β -sheet 2 (β -strands 10–13).

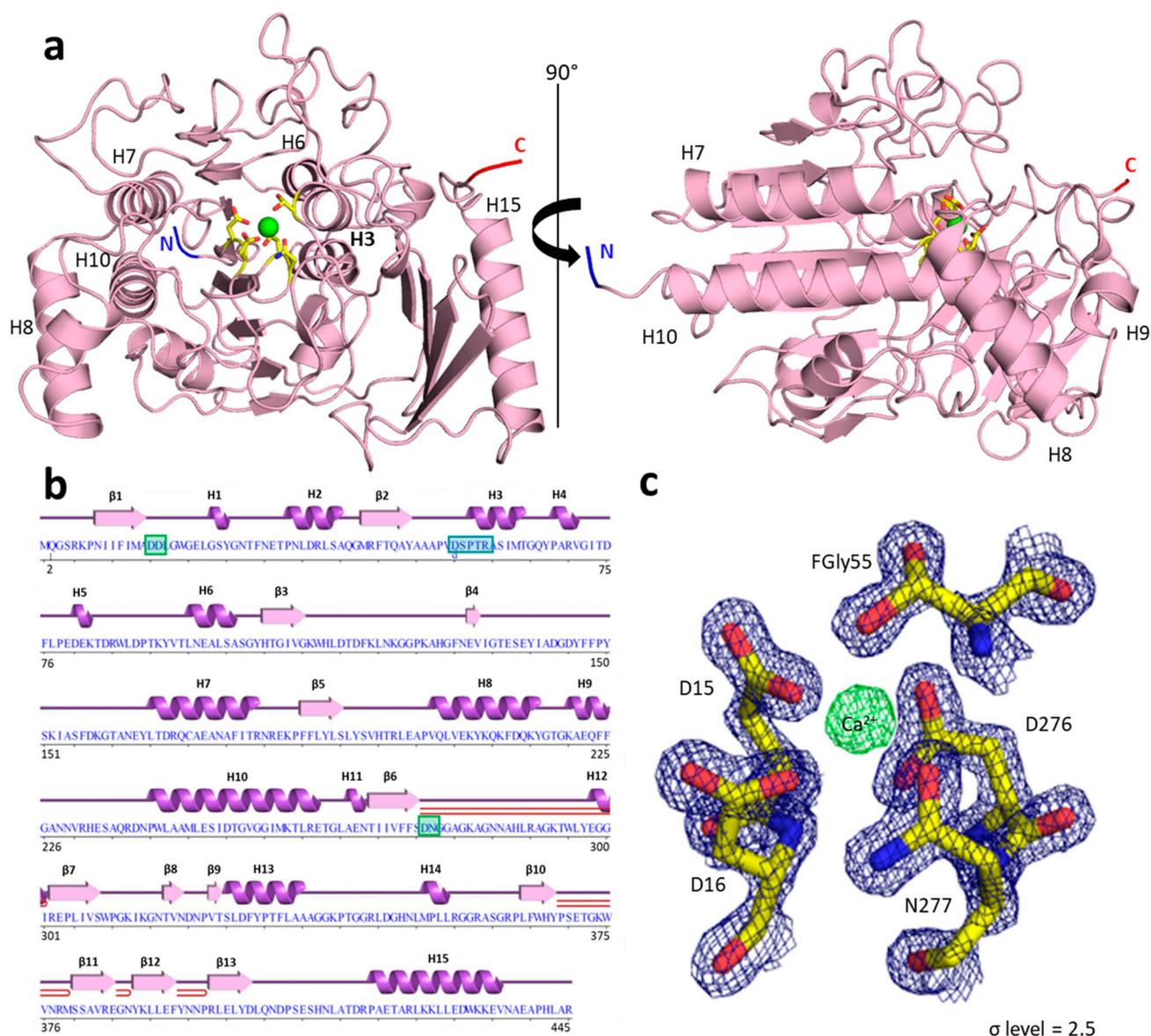


Figure 2. (a) Crystal structure of PyuS calcium (green) bound in the active site (PDB code: 7STT). Residues coordinating the calcium ion are shown in yellow. The N- and C-termini are colored in blue and red, respectively. The formylglycine residue is located in helix 3 (H3). (b) Amino acid sequence and corresponding secondary structural elements of PyuS with labeled α -helices (H) and β -strands (β). The β -hairpins are shown in red. Metal-coordinating residues are highlighted in green. The sulfatase FG motif is highlighted in cyan, and FGly is annotated with a lowercase “d”. (c) Electron density map (contoured at 2.5σ) of active site residues coordinating calcium (D15, D16, FGly55, D276, and N277). In the density, the FGly residue can clearly be seen as a hemiacetal side chain rather than a 2-amino-3-oxopropanoic acid.

The crystal structure of PyuS shows that the active-site cysteine is converted to formylglycine. However, with no sulfate substrate bound in the active site, the expected 2-amino-3-oxopropanoic has been oxidized to a hemiacetal functional group (Figure 2c). This is not an uncommon occurrence among crystal structures of sulfatases with this type of post-translational modification (see PDB codes: 1HDH, 4FDJ, 6G60, 1E33). The aldehyde function of the formylglycine may be stabilized as a hemiacetal functional group, as evidenced by experiments with sodium borohydride and guanidine-denatured sulfatases [24]. In addition, the analysis of the electron density maps, B-factors, and geometry of the metal binding site allowed us to identify Ca^{2+} as being present in the

active site (Figure 2c). Based on our analysis, we assigned a metal occupancy of 60%. The metal binding site was validated with CheckMyMetal server [25]. The calcium ion was further verified by inductively coupled plasma-mass spectrometry (ICP-MS). In a purified sample of PyuS not supplemented with calcium, 165 ppb ($\mu\text{g/L}$) Ca^{2+} was detected, which corresponds to approximately two Ca^{2+} cations per PyuS molecule. The concentration of magnesium in the sample was also determined, and the protein sample contained only 7.3 ppb of Mg^{2+} . These results clearly indicate that the recombinant PyuS sample contains calcium bound in the active site.

2.4. PyuS Oligomerization

The PDBePISA [23] analysis of PyuS crystal structures revealed that PyuS forms homodimers. However, results from size exclusion chromatography (SEC) suggest that only a small population of PyuS dimers may exist in solution. The purification of recombinant PyuS by SEC in Tris-buffered saline (TBS) gave two elution peaks (Figure 3). Proteins from both elution peaks were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions (Figure 3), and both peaks show a protein band at approximately 50 kDa. SEC MW standards were used to calculate the MW of both peaks (Supplementary Figure S1); as expected, the MW of the first elution peak is approximately twice as large as the second elution peak, suggesting that the peaks represent dimers and monomers, respectively.

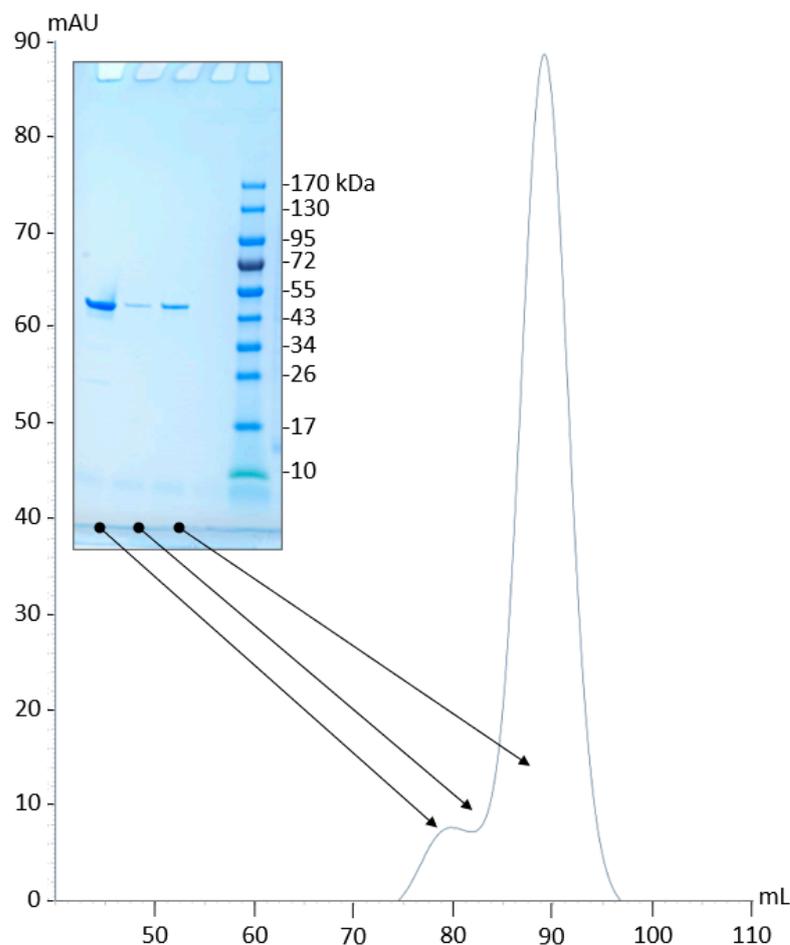


Figure 3. SDS-PAGE gel and size exclusion chromatography purification of recombinant PyuS. For SEC, the absorbance was monitored at 280 nm. The arrows indicate the elution fraction analyzed by SDS-PAGE and the MW of PyuS is 51.3 kDa.

In the 7STT PyuS crystal structure, the interface area of a PyuS homodimer is calculated to be 1023 \AA^2 . This interface area is above the cut-off value of 856 \AA^2 which was established by Ponstingl et al. [26] for differentiating between monomeric and homodimeric proteins in the crystalline state. However, the interface area calculated between monomers in 7STU (bromine) and 7STV (citrate) crystal structures were 876 \AA^2 and 793 \AA^2 , respectively, though the differences in interface areas could be attributed to the crystallization conditions. A PyuS homodimer structure generated by PDBePISA [23] is shown in Figure 4a for PDB code: 7STT. The interface mainly consists of intercalated, charged residues from Helix 5 (residues 77–85), β -hairpin 2 (residues 369–379), and the C-terminal end (residues 439–445) of each monomer in the dimer. In addition, the C-terminal is more disordered in the 7STU and 7STV structures compared to 7STT, and the C-terminal His-tag is not visible in the electron density for any of the structures. For 7STT (Figure 4a), there is also a Na^+ binding site involving residues Asp80 and Thr83. The 7STU (Figure 4b) and 7STV (Figure 4c) structures contain Br^-/Cl^- binding sites (residues Asn377 and Arg 378), respectively, between their monomer interfaces.

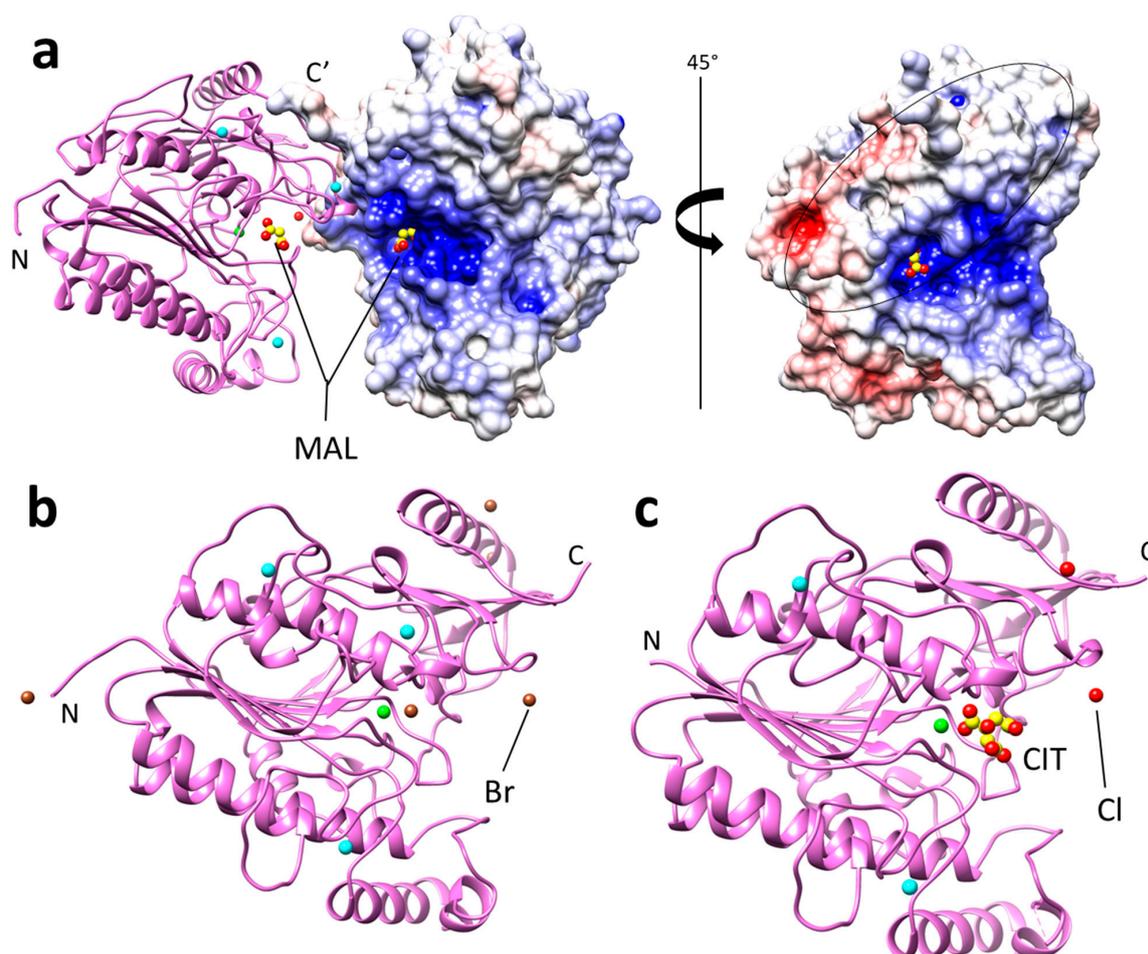


Figure 4. (a) Predicted homodimer of PyuS (PDB code: 7STT) with calcium (green) and malonate (MAL) bound in the active site, as well as sodium (teal) and chlorine (red) ions bound on the protein surface. The dimer interface is mainly formed through electrostatic interactions (black oval). One monomer is shown in cartoon representation (pink) while the other is shown in a surface representation of the electrostatic map (red indicates negative charge, white indicates neutral charge, and blue indicates positive charge). The N indicates the N-terminal of monomer 1, and C' indicates the C-terminal of monomer 2. (b) PyuS crystal structure 7STU with sodium (teal), calcium (green), and bromine (brown) ions bound. (c) PyuS crystal structure 7STV with citrate (CIT) and sodium (teal), calcium (green), and chlorine (red) ions bound.

2.5. PyuS Activity on Sulfatase and Phosphatase Substrates

2.5.1. PyuS Activity on Aromatic Sulfate Esters and Steroid Sulfates

Initially, the small, aromatic, fluorescent substrate 4-methylumbelliferyl sulfate (4MUS) was used to determine the optimal pH range for recombinant PyuS (Figure 5a). PyuS has the highest activity around neutral pH in the range of 6.5–7.5. In addition, PyuS was supplemented with manganese, calcium, magnesium, or zinc, and the activity at pH 6.5 was determined with 4MUS (Figure 5b).

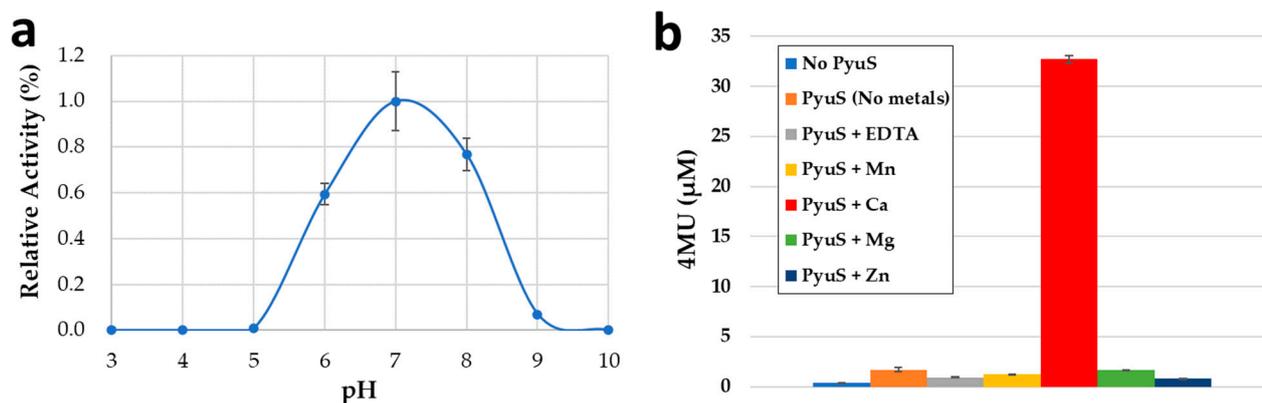


Figure 5. (a) The relative pH optima of PyuS as determined with sulfatase substrate 4MUS. (b) Activity of PyuS on 4MUS in the presence of the different metals. No PyuS, PyuS (No metals), and EDTA were used as controls. Error bars represent biological replicates.

Based on the metal supplement assay with substrate 4MUS, calcium is the preferred co-factor for PyuS, as expected based on data from the crystal structures (Figure 2c). The PyuS sample (no metals) in Figure 5b contains 4MUS substrate but does not contain any supplemented co-factor; comparing only PyuS to PyuS + Ca²⁺, it is apparent that adding additional calcium to the reaction will increase the activity of PyuS and provides evidence that calcium is the preferred co-factor for the catalysis of sulfate esters. In addition, the metal chelator ethylenediaminetetraacetic acid (EDTA) was used as a negative control for the 4MUS reactions (Figure 5b), and as expected, PyuS had lower activity in the presence of the EDTA when compared to the PyuS (no metals) sample.

The dimeric and monomeric elution samples from SEC (Figure 3) were concentrated separately, and their activities on 4MUS were compared at pH 6.5 (Figure 6a). The dimeric PyuS has an activity of 2.12 nmol/min/mg on 4MUS, whereas the monomeric PyuS has an activity of 2.50 nmol/min/mg. These results indicate that the oligomerization state of PyuS does not have a substantial impact on 4MUS activity. The activity of the monomer on 4MUS was investigated further with Michaelis–Menten kinetics (Figure 6b,c). Because the enzyme appears to exist as both dimer and monomer, kinetic data were modeled by using both the Michaelis–Menten equation and the Hill equation. The Michaelis constant (K_m) was calculated to be 1.89 mM, the velocity max (V_{max}) at 1.16 $\mu\text{M}/\text{min}$, turnover number (k_{cat}) at 0.59 min^{-1} , and a Hill coefficient of approximately 1.0 was determined, suggesting that there is no obvious cooperativity between PyuS active sites.

PyuS activity with two other general arylsulfatase substrates, *para*-nitrocatechol sulfate (pNCS) and 5-bromo-4-chloro-3-indoxyl sulfate (X-Sulf), was also determined (Figure 6d). Compared to 4MUS hydrolysis, the activity of PyuS on pNCS is remarkably slow; it should be noted that the 4MUS assay was incubated for 21 h (Figure 6a), whereas the pNCS assay (Figure 6d) was incubated for over 48 h, resulting in low product formation (see Materials and Methods for more details). PyuS did not have any significantly apparent activity on X-Sulf (Figure 6d), another substrate which has been used for arylsulfatase screening [27].

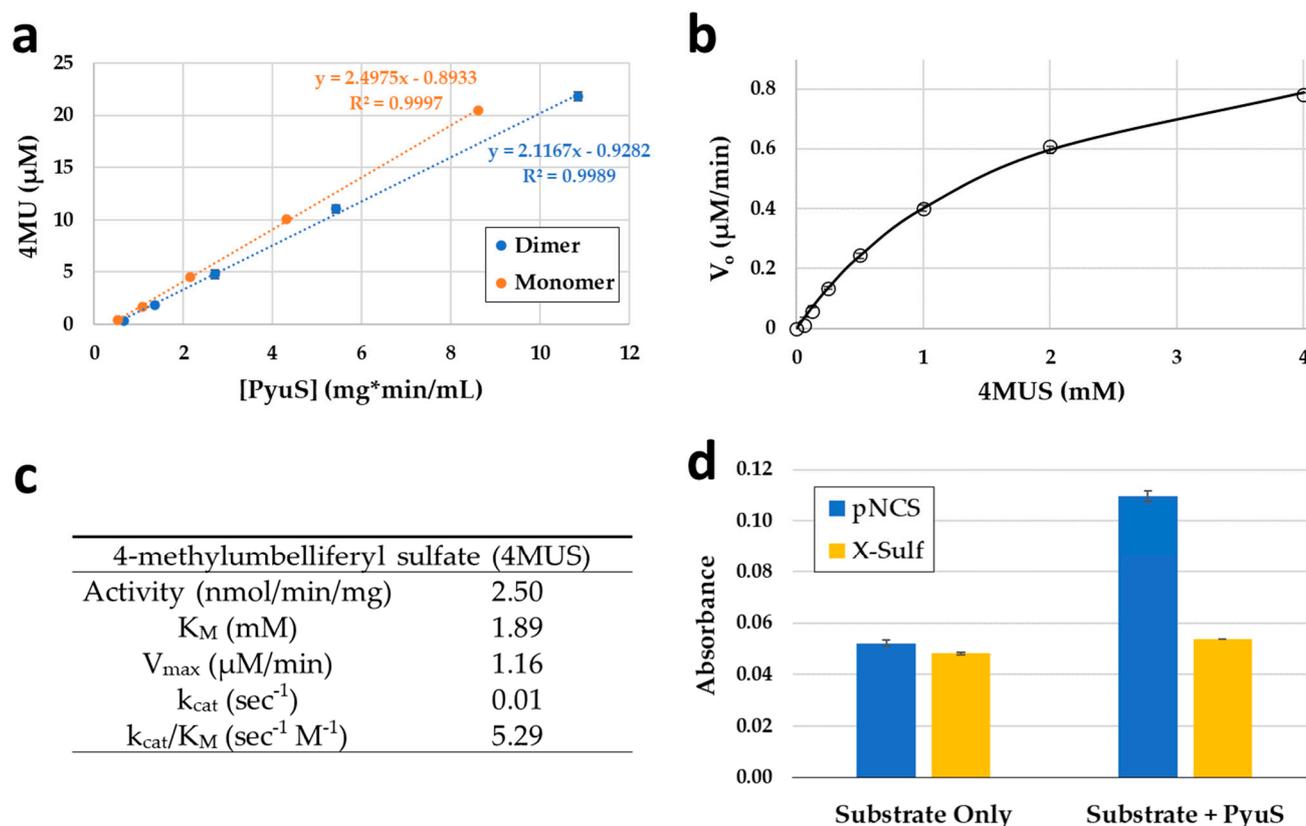


Figure 6. (a) Activity curves of dimeric and monomeric recombinant PyuS on substrate 4MU. (b) Enzyme kinetics curve for 4MUS fit with the Michaelis–Menten equation. (c) PyuS kinetic parameters on 4MUS substrate. (d) Recombinant PyuS activity on arylsulfatase substrates pNCS and X-Sulf at pH 6.5. The pNCS and X-Sulf products were measured at 515 nm and 650 nm, respectively. Assays were performed with biological replicates.

Some arylsulfatases, such as PaeS from *P. aeruginosa*, are capable of hydrolyzing steroid sulfates [4,5]. The activity of PyuS on the steroid sulfates cortisol 21-sulfate (CS), 17α -estradiol sulfate (α ES), and boldenone sulfate (BS) was measured by high-pressure liquid chromatography (HPLC). However, PyuS was not able to hydrolyze any of the tested steroid substrates (data not shown).

2.5.2. PyuS Activity on Phosphatase Substrates

Due to the evolutionary relationship of sulfatase and alkaline phosphatase enzyme families [28] and the promiscuity of both activities, the phosphatase substrates *para*-nitrophenyl phosphate (pNPP) and 5-bromo-4-chloro-3-indoxyl phosphate (X-Phos) were incubated with PyuS to monitor phosphatase activity. Interestingly, PyuS is able to hydrolyze both pNPP and X-Phos (Figure 7).

2.6. PyuS Stability

The melting temperature (T_m) of recombinant PyuS was determined by measuring the change in intrinsic tryptophan fluorescence emission wavelength due to changes in the polarity of the surrounding environment, and this value was recorded as the barycentric mean (BCM) (Figure 8). Stability was measured in TBS, TBS with calcium, TBS with potassium sulfate, sodium acetate pH 6.5, and sodium acetate with calcium. The condition of sodium acetate pH 6.5 supplemented with calcium resulted in the highest T_m of approximately 48°C , whereas all other conditions had a T_m of approximately 45°C . Sodium acetate was tested because it was the buffer of choice for hydrolysis assays.

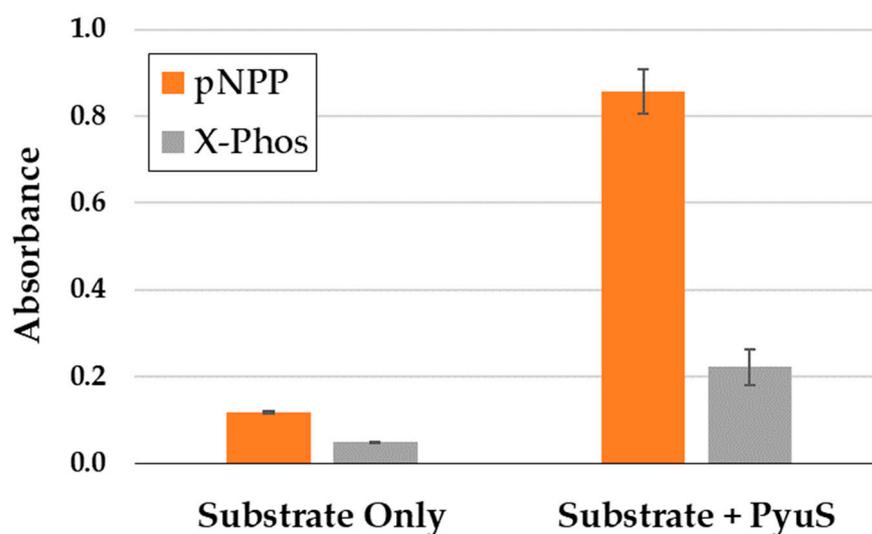


Figure 7. Activity of PyuS on phosphatase substrates pNPP and X-Phos. Products of pNPP and X-Phos were measured at 405 nm and 650 nm, respectively. Assays were performed with biological replicates.

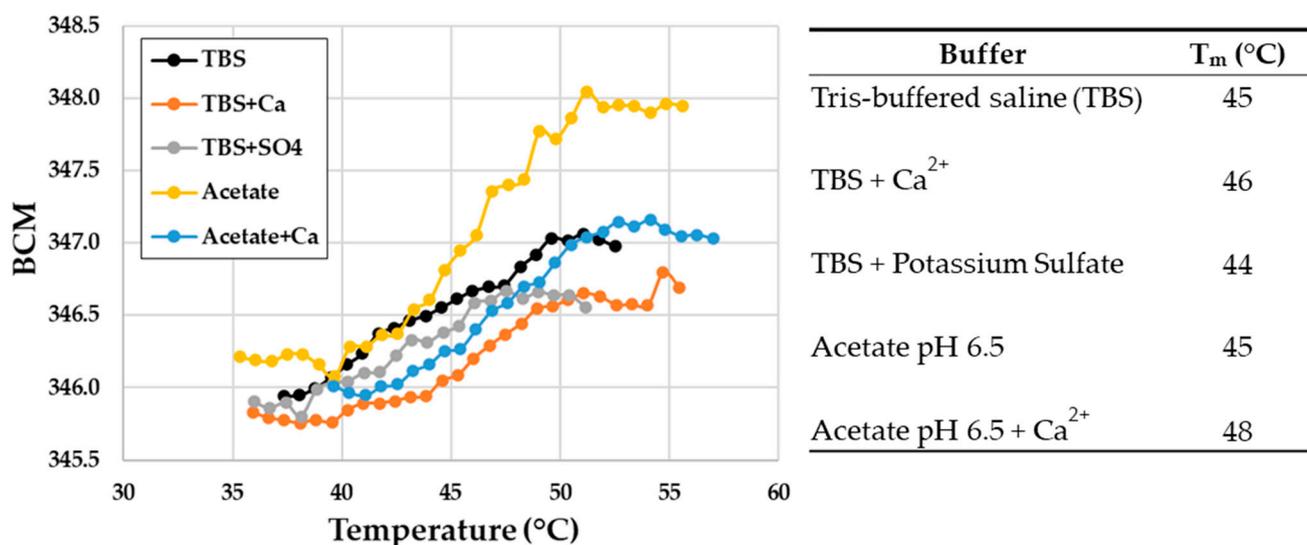


Figure 8. Melting curves (T_m) for recombinant PyuS in various buffers. Samples were biological replicates.

3. Discussion

We showed that the putative sulfatase PyuS from *P. yulinensis* contains the sulfatase motif CXPXR and that the cysteine is post-translationally converted into a formylglycine when co-expressed with an FGE in *E. coli*. However, in the electron density of the crystal structures, the formylglycine is observed as a hemi-acetal group, and it is unclear whether this is the naturally preferred resting state of the residue or just an artifact of purification or crystallization. PyuS is most active near neutral pH 7 on substrate 4MUS, uses calcium as a cofactor, and hydrolyzes the sulfatase substrates 4MUS and pNCS, as well as phosphatase substrates pNPP and X-Phos. However, it has no detectable activity on steroid sulfates α ES, CS, or BS. Calcium was the co-factor of choice for 4MUS hydrolysis, but it is possible that PyuS may utilize a different co-factor, such as magnesium, for the hydrolysis of other substrates, and this concept would need to be explored further.

Based on size-exclusion chromatography, approximately 10% of soluble PyuS will form homodimers (Figure 3). Dimers are also seen in the crystal structures, and PDBePISA [23]

shows that the interface area between monomers in the crystal structure exceeds the threshold to distinguish between real dimers and experimental artifacts. Exploring the interface area further reveals that mostly electrostatic interactions permit dimerization, and there are no inter- or intramolecular disulfide bonds (other than the Cys in CXPXR motif, PyuS only has one cysteine residue which is not solvent exposed). In all of the PyuS crystal structures (Figure 4), there is at least one type of ion (sodium, bromine, or chloride) present in the putative dimer interface, and these ions may help facilitate dimer formation.

It should be reiterated that the C-terminal electron density was disordered for all structures, where 7STU and 7STV C-termini had more disorder than that in 7STT. The C-terminal ends in the putative dimer (Figure 4a) are only approximately 20 Å apart as measured from the C α 's of Arg445 from each monomer. At this proximity, it is plausible that the His-tag from each monomer could interact with one another [29,30], and the dimers we observe during purification and from crystallization could be a consequence of this interaction. However, it should also be considered that the His-tag could be preventing oligomerization, and if the His-tag was moved to the N-terminal or if a different purification tag was used altogether, it is possible that PyuS oligomerization would be more prominent.

Although a low percentage of soluble PyuS is in an oligomeric state, we cannot rule out the possibility that this may change in the presence of substrate or by increasing the concentration of PyuS. The activity of the dimeric and monomeric PyuS (Figure 4a) was compared on 4MUS, and there was a minimal difference in activity observed (Figure 6a). In the homodimer, the active sites are approximately 33 Å apart as measured between active site calcium ions in PyMOL [31]. The cooperativity could be further investigated by modifying residues in Helix 5 (residues 77–85) and β -hairpin 2 (residues 369–379, Figure 2b) as these structural elements are in close proximity to the active site of the opposing monomer in the homodimer.

Because there are generally low sequence identities among sulfatases, we used several homologous protein structure-search programs such as PDBeFold [32,33], Dali [34], PDBsum [35], CATH [36,37], and PDBeMotif [38] to find homologous structures to PyuS (Table 1).

Table 1. Protein Data Bank (PDB) sulfatase structures homologous to PyuS. RMSD is the root mean square deviation from C α alignment; C α is the number of alpha carbons aligned; and sequence identity refers to primary amino acid sequence alignment. RMSD values were calculated by aligning C α in PyMOL [31]. Sequence identity was calculated with CLUSTAL W [39]. Signal peptides or putative signal peptides were removed for all alignments.

PDB Code	Protein	RMSD (Å)	C α Aligned	Sequence Identity to PyuS (%)
1FSU	N-acetylgalactosamine-4-sulfatase	1.0	265	18.4
1HDH	Arylsulfatase	1.7	260	15.3
1P49	Sterylsulfatase	0.9	251	20.0
6USS	Sulfatase (not classified)	0.8	275	24.9
6UST	N-acetylglucosamine-6-sulfatase	0.8	317	29.8
5G2V	N-acetylglucosamine-6-sulfatase	1.7	251	19.8
4FDJ	N-acetylglucosamine-6-sulfatase	0.9	249	23.3
6B0J	Iota-carrageenan sulfatase	0.9	277	23.1

As expected, sulfatases were the most prominent homologous proteins found when searching with the PyuS structure, but there were a few alkaline phosphatase structures in the list as well which had lower homology than the sulfatases (Table 1). Interestingly, the sulfatases with high homology to PyuS had variable functions which include N-acetylgalactosamine-4-sulfatases, N-acetylglucosamine-6-sulfatases, arylsulfatases, steryl-sulfatases, and Iota-carrageenan sulfatases (Table 1). The sulfatase structures in Table 1 were aligned with PyuS in PyMOL [31], and the root mean square deviation (RMSD) was generally close to 1.0 Å or better when aligning the structures by alpha carbons (C α).

The active site of PyuS contains His113, His195, and Lys293 which are completely conserved among the sulfatases listed in Table 1. In addition, the PyuS active site contains

Phe76 that is not conserved among sulfatases with known crystal structures in the PDB. For example, Phe76 is replaced by tyrosine in the *H. sapiens* N-acetylgalactosamine-6-sulfatase (PDB code: 4FDJ) or tryptophan in the *Bacteroides fragilis* sulfatase (PDB code: 6USS). Our structural studies indicate that negatively charged entities such as malonate, bromine, and citrate are favored ligands for binding to the PyuS active site, and these ligands are in the vicinity of active site residues His195 and Lys293. This is not surprising as they mimic sulfate or the phosphate moieties of sulfatase or phosphatase substrates, respectively, and the active site of PyuS has an overall positive charge (Figure 4a).

A C α alignment of all the sulfatase structures in Table 1 shows that the sulfatase active site is comprised of several unstructured loops, similarly to PyuS (Figure 2a), and it is not far-fetched to believe these loops influence substrate specificity and the pH optima of the enzymes. Based on our kinetic studies, the sequence alignments, and structure alignments, it is plausible that PyuS is actually an N-acetylgalactosamine-4-sulfatase or N-acetylglucosamine-6-sulfatase instead of an arylsulfatase. If that is not true, then PyuS may be involved in the metabolic pathways of *P. yulinensis*, and scavenging sulfate [40] or phosphate-containing [41] compounds may be its primary role.

4. Materials and Methods

4.1. Gene Synthesis, Protein Expression, and Purification

The amino acid sequence of PyuS (Accession No. WP_107215127) was found by using the protein Basic Local Alignment Search Tool (BLASTp, [17]) with sequences from *P. aeruginosa* (Accession No. WP_003106692), *H. pomatia* (Accession No. AF109924), or *H. sapiens* (Accession No. NP_000037, AAA60596). A codon-optimized nucleic acid sequence for PyuS was synthesized by BioBasic (Toronto, Canada) for *E. coli* expression. The N-terminal region of native PyuS (amino acid residues 2–24) was predicted to be a signal peptide by the SignalP server [18], and these residues were removed from the DNA sequence for expression. A C-terminal His6-tag (residues GSHHHHHH) was added for IMAC purification, and the final molecular weight of the recombinant PyuS construct was 51,320 Da (459 amino acids). In addition, the FGE gene from *M. tuberculosis* was ordered from BioBasic (Toronto, Canada) for co-expression with PyuS in *E. coli* [15,20]. Both genes were cloned into pCRS195, a proprietary plasmid synthesized (by Life Technologies Corporation, Carlsbad, CA) for use by Integrated Micro-Chromatography Systems (IMCS). Genes were incorporated into pCRS195 with Bsu36I, BamHI, AvrII, and MfeI restriction sites; restriction enzymes were purchased from New England Biolabs (NEB, Ipswich, MA, USA). In pCRS195, genes are expressed under the control of a repressible T7 promoter. The resulting plasmid containing both FGE and PyuS genes was named pCRS268.

E. coli T7 expressing cells from NEB (Ipswich, MA, USA) were transformed with pCRS268 and grown on LB-kanamycin (KAN, 50 μ g/mL) agar plates. A starter culture was made from a single colony in LB-KAN, and this culture was grown overnight at 37 °C. Starter culture was added to 700 mL of LB-KAN, grown to O.D. 0.8 at 37 °C, and a working concentration of 50 μ M copper sulfate was added [42]. The culture was cooled at 16 °C for 10 min and then induced at 16 °C with 0.1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) for 16 h. Cell pellets were harvested and frozen at –80 °C until further needed.

Cells were lysed with Bacterial Protein Extraction Reagent (B-PER™, ThermoFisher Scientific, Waltham, MA, USA) following the manufacturer's protocol. Soluble cell lysate was purified with a 1.0 mL nickel-IMAC column attached to an ÄKTA pure protein purification system (Cytiva, Marlborough, MA, USA) at 4 °C. The loading buffer used was 10 mM Tris, 100 mM sodium chloride, 10 mM imidazole, pH 8. The elution buffer was the same but with the addition of 500 mM imidazole. The IMAC column was washed with 5%, 10%, 40%, and 100% of elution buffer, and PyuS was found in both the 10% and 40% elution fractions. Protein was concentrated with an Amicon concentrator (EMD Millipore, Burlington, MA, USA) with a 30 kDa cut-off membrane. Protein was purified further by SEC with a Cytiva HiLoad 16/600 superdex 200 pg column, using 10 mM Tris, 150 mM sodium chloride, pH 8.0 as the running buffer at a flow rate of 1.2 mL/min and 4 °C. Protein purity

was analyzed by using pre-cast 4%–20% gradient SDS-PAGE gels from BioRad (Hercules, CA, USA). Protein concentration was measured by determining the absorbance at 280 nm and using the ExPASy [21] predicted extinction coefficient of $77,810 \text{ M}^{-1} \text{ cm}^{-1}$. PyuS has an isoelectric point of 8.24. The recombinant PyuS yield from a 700 mL LB culture was approximately 5 mg.

4.2. Metal Analysis

The presence of calcium and magnesium in the protein was evaluated with inductively coupled plasma–mass spectrometry (ICP-MS). Nitric acid was added to PyuS (0.5 mg protein in 500 μL of 10 mM Tris, 150 mM NaCl, pH 8) to a final concentration of 2% HNO_3 before analysis. A Finnigan ELEMENT XR double focusing magnetic sector field (SF) ICP-MS was used for the analysis with Rh as an internal standard. A 0.2 mL/min Micromist U-series nebulizer (GE, Port Melbourne, Victoria, Australia), quartz torch, and injector (Thermo Fisher Scientific, Waltham, MA, USA) were used for sample introduction. Five-point calibration curves were used in a calibration range from 10 to 600 ppb.

4.3. Protein Crystallization, Data Collection, and Refinement

Protein crystals were obtained via the vapor diffusion technique. The 96-well crystallization plates (Hampton Research, Aliso Viejo, CA, USA) were left at room temperature (23 °C) for all experiments. Crystallization drops were set with 1.0 μL of protein solution with protein concentrations varying from 4 to 6 mg/mL in 10 mM Tris, 150 mM NaCl, pH 8.0 buffer. The protein was mixed with 1.0 μL of crystallization solution in the wells. Protein was also incubated on ice for 30 min with sulfatase irreversible inhibitor STX64 (Axon Medchem, Reston, VA, USA) [43,44] with a final inhibitor concentration of 1.0 mM before setting drops. However, the inhibitor did not solubilize well in the protein mixture and precipitated out of solution in the crystallization drops. No crystals were found to contain the STX64 inhibitor.

Protein crystals were grown in wells containing 2.4 M sodium malonate pH 7.0 (PDB code: 7STT), 0.15 M potassium bromide and 30% *w/v* polyethylene glycol monomethyl ether 2000 (PDB code: 7STU), or 0.1 M sodium citrate pH 5.5 and 20% *w/v* polyethylene glycol 3000 (PDB code: 7STV). Crystals were cryocooled in liquid nitrogen prior to data collection. All X-ray diffraction experiments were performed at the South East Regional Collaborative Access Team beamline 22ID at the Advanced Photon Source (Argonne National Laboratory, Lemont, IL, USA). HKL-2000 was used for data processing [45]. Data collection statistics for the three structures are reported in Table 2.

Table 2. Data collection and refinement statistics for PyuS crystal structures. Values in parentheses are for the highest resolution shell. RMSD—root mean square deviation.

Structure	PyuS (Malonate)	PyuS (Bromine)	PyuS (Citrate)
PDB accession code	7STT	7STU	7STV
Data collection			
Diffraction source	APS, 22ID	APS, 22ID	APS, 22ID
Wavelength (Å)	1.0	1.0	1.0
Space group	P3 ₂ 21	P3 ₂ 21	P3 ₂ 21
a, b, c (Å)	83.6, 83.6, 116.1	83.6, 83.6, 115.8	82.7, 82.7, 115.3
Resolution range (Å)	40.0–1.60 (1.63–1.60)	40.0–2.23 (2.27–2.23)	40.0–2.34 (2.38–2.34)
No. of unique reflections	59,449 (2069)	23,149 (1144)	19,275 (936)
Completeness (%)	95.7 (67.5)	99.2 (99.6)	98.8 (98.7)
Redundancy	10.6 (5.8)	9.1 (6.2)	9.2 (6.7)
$\langle I/\sigma(I) \rangle$	35.3 (2)	15.1 (2)	18.5 (2)
R _{meas}	0.079 (0.487)	0.172 (0.858)	0.148 (0.525)
R _{p.i.m.}	0.024 (0.186)	0.057 (0.339)	0.048 (0.175)
CC1/2	0.996 (0.859)	0.997 (0.738)	0.998 (0.948)

Table 2. Cont.

Structure	PyuS (Malonate)	PyuS (Bromine)	PyuS (Citrate)
Refinement			
Resolution range (Å)	39.37–1.60 (1.63–1.60)	39.22–2.23 (2.29–2.23)	34.23–2.35 (2.41–2.35)
Completeness (%)	95.5 (68.4)	99.2 (99.3)	98.4 (93.3)
No. of reflections, working set	56312	21815	18285
No. of reflections, test set	3033	1205	943
Final Rcryst	0.148 (0.276)	0.174 (0.282)	0.172 (0.286)
Final Rfree	0.178 (0.270)	0.225 (0.319)	0.230 (0.301)
RMSD bonds (Å)	0.014	0.007	0.005
RMSD angles (°)	1.8	1.4	1.3
Ramachandran plot			
Allowed regions (%)	100	100	100
Favored regions (%)	97.1	97.1	97.2

4.4. Structure Determination, Refinement, and Validation

The structures were solved by molecular replacement using MOLREP [46] and HKL-3000 [47]. The search model for 7STT was a predicted structure from SWISS-Model [48–50] using a structure of an N-acetylglucosamine-6-sulfatase (PDB code: 6UST) with 38.7% sequence identity to PyuS. The 7STT structure was then used as the search model for the other two structures (7STU and 7STV). The structures were refined with REFMAC [51] and COOT [22]. The TLS Motion Determination server [52] was used to split the single protein chain into fragments in the final stages of refinement for each of the three structures. MOLPROBITY [53] and COOT were used for structure validation. The identities and locations of metals in the structures were validated with the CheckMyMetal (CMM) server [25,54,55]. Table 1 shows the final refinement and validation statistics. The three structures were deposited in the PDB with accession codes 7STT, 7STU, and 7STV.

4.5. Protein Modeling and Structure Analysis

The predicted crystal structure of HpoS was generated by using SWISS-Model [48–50]. Protein structures were analyzed with PDBEPIA [23], PDBeFold [32,33], Dali [34], PDBsum [35], CATH [36,37], and PDBeMotif [38]. Protein crystal structures were modeled in COOT [56], PyMOL [31] or Chimera [57]. Electrostatic maps were generated or modified with APBS [58] or CCP4 package suite [59]. Protein sequences were aligned with the program CLUSTAL W [39].

4.6. Hydrolysis Assays with 4MUS, pNCS, and X-Sulf

All chemical structures of substrates can be found in Supplementary Figure S2. Substrates were purchased from Sigma-Aldrich (St. Louis, MO, USA), GoldBiotechnology (St. Louis, MO, USA), or Fisher Scientific (Waltham, MA, USA). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) or ThermoFisher (Waltham, MA, USA). The pH optimum of PyuS was determined using a triple buffer system of acetic acid, MES, and Tris as described by Ellis and Morrison [60] with 4-methylumbelliferyl sulfate at a working concentration of 1.0 mM, calcium chloride at 1.0 mM, and PyuS at approximately 0.05 mg/mL at room temperature (RT) for 2 h. The metal co-factor assay used a working concentration of 1.0 mM manganese chloride, calcium chloride, magnesium chloride, zinc acetate, or EDTA in 100 mM sodium acetate, pH 6.5, with 0.2 mg/mL PyuS and 1.0 mM 4MUS for 3 h at RT. The 4MUS standard assay was performed with a working concentration of 1.0 mM 4MUS and 0.5 mM calcium chloride in 50 mM sodium acetate, pH 6.5, for 21 h at RT. The *para*-nitrocatechol sulfate (pNCS) and 5-bromo-4-chloro-3-indoxyl sulfate (X-Sulf) assays were performed at 1.0 mM and 0.3 mg/mL concentrations, respectively, in 100 mM sodium acetate pH 6.5, 0.5 mM calcium chloride, and 0.3 mg/mL PyuS with a 60 h incubation at RT. All 4MUS, pNCS, and X-Sulf reactions were stopped with 200 mM glycine, pH

10.4. Excitation and emission wavelengths used for the detection of 4-methylumbelliferone (4MU) were 360 nm and 449 nm, respectively. The absorbance of products from pNCS and X-Sulf hydrolysis were detected at 515 nm and 650 nm, respectively. All assay products were detected in a Spectramax M5 (Molecular Devices, San Jose, CA, USA). Hydrolysis assays were performed in duplicate.

For Michaelis–Menten kinetics, the highest working concentration was 4.0 mM 4MUS. The 4MUS was serially diluted 1:1 in 25 mM Bis-Tris, pH 6.5, and 1.0 mM calcium chloride, to 0.06 mM 4MUS which served as the lowest concentration. The working concentration of PyuS was 0.1 mg/mL, and reactions were 50 μ L total and proceeded for up to 1.5 h at room temperature. Reactions were stopped with the addition of 150 μ L 200 mM glycine, pH 10.4. All reactions were performed in duplicate except 4.0 mM and 0.06 mM 4MUS, which were single reactions. Kinetic parameters were determined with the Michaelis–Menten or Hill equation and by using the Microsoft Excel Solver add-on.

4.7. HPLC with Steroid Sulfates

Steroids were purchased from Toronto Research Chemicals (Toronto, ON, Canada) or Cerilliant (Round Rock, TX, USA). For assay setup, 0.5 mg/mL PyuS in 100 mM sodium acetate pH 6.5 and 1.0 mM calcium chloride was mixed with 0.1 mg/mL 17 α -estradiol sulfate or 0.01 mg/mL cortisol 21-sulfate or boldenone sulfate at RT for 24 h. Samples were stopped with 41% MeOH, 11% acetonitrile, 10 mM potassium phosphate pH 7.6, and 0.1% formic acid (*v/v*). Ultra-performance liquid chromatography was performed on a Thermo Scientific™ Vanquish™ UHPLC system using a Thermo Scientific™ Accucore™ biphenyl reversed phase 80 Å column (2.1 mm \times 50 mm, 2.6 μ m). The column was heated to 35 °C and used with a flow rate of 0.5 mL/min. Mobile phase used was HPLC running buffer: 41% methanol, 11% acetonitrile, 10 mM potassium phosphate, pH 7.6. The absorbance was measured at 250 nm for CS and BS, and 280 nm for α ES. PyuS had no detectable activity on these steroids. Samples were performed in duplicates.

4.8. Hydrolysis Assays with pNPP and X-Phos

The *para*-nitrophenyl phosphate was purchased from Millipore Sigma (St. Louis, MO, USA), and X-Phos was purchased from Fisher Scientific (Waltham, MA, USA). The pNPP and X-Phos were at 1.0 mM and 0.4 mg/mL, respectively, with 100 mM sodium acetate pH 6.5, 0.5 mM calcium chloride, and 0.3 mg/mL PyuS incubated at RT for 60 h. Reactions were stopped with 200 mM glycine, pH 10.4. Products of pNPP and X-Phos were read at wavelengths of 405 nm and 650 nm, respectively, in a Spectramax M5 (Molecular Devices, San Jose, CA, USA). Samples were performed in duplicates.

4.9. Melting Temperature of PyuS

The melting temperature for PyuS was measured by using an UNcle instrument from UNChained Labs (Woburn, MA, USA). Protein was used at a working concentration of 0.8 mg/mL, calcium chloride at 1.0 mM, potassium sulfate at 1.0 mM, TBS at 10 mM Tris, 150 mM sodium chloride (pH 8.0), and sodium acetate at 100 mM where applicable. Samples were heated for 180 s at different temperatures in 1.0 °C increments, ranging from 30 °C to 95 °C. The change in intrinsic tryptophan fluorescence emission wavelength was measured at each temperature and recorded as BCM. The melting temperature was calculated from the inflection point of the first derivative. Samples were in duplicates and the standard deviation for all samples was less than zero.

Supplementary Materials: The following are available online, Figure S1: SEC purification of PyuS. Figure S2: Chemical structure of substrates characterized with PyuS.

Author Contributions: Conceptualization, C.R.S., J.J.T. and L.A.L.; purification and kinetic characterization, C.R.S. and L.L.G.; methodology, C.R.S. and J.J.T.; software, C.R.S., L.L.G., A.O. and M.C.; validation, A.O. and M.C.; formal analysis, C.R.S., A.O. and M.C.; investigation, C.R.S., L.L.G., J.J.T. and L.A.L.; resources, L.A.L. and M.C.; data curation, C.R.S.; writing—original draft preparation, C.R.S., A.O. and

M.C.; writing—review and editing, C.R.S., L.L.G., A.O., J.J.T., M.C. and L.A.L.; visualization, C.R.S.; supervision, C.R.S. and J.J.T.; project administration, L.A.L. and M.C.; funding acquisition, L.A.L. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: Structural results shown in this report were derived from data collected at Southeast Regional Collaborative Access Team (SER-CAT; 22 ID) beamline at the Advanced Photon Source, Argonne National Laboratory. Supporting institutions may be found at https://www.ser.aps.anl.gov/www.ser-cat.org/images/SER-CAT18_Member_list.jpg. The use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract Nos. DE-AC02-06CH11357 and W-31-109-Eng-38. This work was partially supported by an ASPIRE III grant from the Office of the Vice President of Research at the University of South Carolina.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All crystal structures are available from the Protein Data Bank (PDB) with accession codes 7STT, 7STU, and 7STV.

Acknowledgments: We would like to thank Lijian He at the University of South Carolina for assistance with ICP-MS analysis.

Conflicts of Interest: The authors declare no conflict of interest. The IMCS funder is a lab that researches and develops drug tests for the characterization of drugs of abuse including sulfated metabolites.

Sample Availability: Raw diffraction images are available from the authors upon request.

References

1. Varin, L.; DeLuca, V.; Ibrahim, R.K.; Brisson, N. Molecular Characterization of Two Plant Flavonol Sulfotransferases. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1286–1290. [[CrossRef](#)] [[PubMed](#)]
2. Das, S.; Sreevidya, V.S.; Udvardia, A.J.; Gyaneshwar, P. Dopamine-Induced Sulfatase and Its Regulator Are Required for Salmonella Enterica Serovar Typhimurium Pathogenesis. *Microbiology* **2019**, *165*, 302–310. [[CrossRef](#)] [[PubMed](#)]
3. Reed, M.J.; Purohit, A.; Woo, L.W.L.; Newman, S.P.; Potter, B.V.L. Steroid Sulfatase: Molecular Biology, Regulation, and Inhibition. *Endocr. Rev.* **2005**, *26*, 171–202. [[CrossRef](#)] [[PubMed](#)]
4. Stevenson, B.J.; Waller, C.C.; Ma, P.; Li, K.; Cawley, A.T.; Ollis, D.L.; McLeod, M.D. Pseudomonas Aeruginosa Arylsulfatase: A Purified Enzyme for the Mild Hydrolysis of Steroid Sulfates: Steroid Sulfate Hydrolysis by Pseudomonas Aeruginosa Arylsulfatase. *Drug Test. Anal.* **2015**, *7*, 903–911. [[CrossRef](#)]
5. Iannone, M.; Botrè, F.; Martinez-Brito, D.; Matteucci, R.; de la Torre, X. Development and Application of Analytical Procedures for the GC–MS/MS Analysis of the Sulfates Metabolites of Anabolic Androgenic Steroids: The Pivotal Role of Chemical Hydrolysis. *J. Chromatogr. B* **2020**, *1155*, 122280. [[CrossRef](#)]
6. Sasisekharan, R.; Raman, R.; Prabhakar, V. Glycomics Approach to Structure-Function Relationships of Glycosaminoglycans. *Annu. Rev. Biomed. Eng.* **2006**, *8*, 181–231. [[CrossRef](#)]
7. Diez-Roux, G.; Ballabio, A. Sulfatases and human disease. *Annu. Rev. Genom. Hum. Genet.* **2005**, *6*, 355–379. [[CrossRef](#)]
8. Dierks, T.; Schmidt, B.; Borissenko, L.V.; Peng, J.; Preusser, A.; Mariappan, M.; von Figura, K. Multiple Sulfatase Deficiency Is Caused by Mutations in the Gene Encoding the Human C α -Formylglycine Generating Enzyme. *Cell* **2003**, *113*, 435–444. [[CrossRef](#)]
9. Barbeyron, T.; Brillet-Guéguen, L.; Carré, W.; Carrière, C.; Caron, C.; Czjzek, M.; Hoebeke, M.; Michel, G. Matching the Diversity of Sulfated Biomolecules: Creation of a Classification Database for Sulfatases Reflecting Their Substrate Specificity. *PLoS ONE* **2016**, *11*, e0164846. [[CrossRef](#)]
10. The UniProt Consortium. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2017**, *45*, D158–D169. [[CrossRef](#)]
11. Knaust, A.; Schmidt, B.; Dierks, T.; von Bülow, R.; von Figura, K. Residues Critical for Formylglycine Formation and/or Catalytic Activity of Arylsulfatase, A. *Biochemistry* **1998**, *37*, 13941–13946. [[CrossRef](#)]
12. Dierks, T. Sequence Determinants Directing Conversion of Cysteine to Formylglycine in Eukaryotic Sulfatases. *EMBO J.* **1999**, *18*, 2084–2091. [[CrossRef](#)] [[PubMed](#)]
13. Berteau, O.; Guillot, A.; Benjdia, A.; Rabot, S. A New Type of Bacterial Sulfatase Reveals a Novel Maturation Pathway in Prokaryotes. *J. Biol. Chem.* **2006**, *281*, 22464–22470. [[CrossRef](#)] [[PubMed](#)]
14. Dierks, T.; Schmidt, B.; von Figura, K. Conversion of Cysteine to Formylglycine: A Protein Modification in the Endoplasmic Reticulum. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11963–11968. [[CrossRef](#)] [[PubMed](#)]
15. Carlson, B.L.; Ballister, E.R.; Skordalakes, E.; King, D.S.; Breidenbach, M.A.; Gilmore, S.A.; Berger, J.M.; Bertozzi, C.R. Function and Structure of a Prokaryotic Formylglycine-Generating Enzyme. *J. Biol. Chem.* **2008**, *283*, 20117–20125. [[CrossRef](#)]
16. Wei, Y.; Wang, B.; Zhang, L.; Zhang, J.; Chen, S. Pedobacter Yulinensis Sp. Nov., Isolated from Sandy Soil, and Emended Description of the Genus Pedobacter. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 2523–2529. [[CrossRef](#)]

17. 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](#)]
18. Nielsen, H. Predicting Secretory Proteins with SignalP. In *Protein Function Prediction*; Kihara, D., Ed.; Methods in Molecular Biology; Springer: New York, NY, USA, 2017; Volume 1611, pp. 59–73, ISBN 978-1-4939-7013-1.
19. Nielsen, H.; Tsirigos, K.D.; Brunak, S.; von Heijne, G. A Brief History of Protein Sorting Prediction. *Protein J.* **2019**, *38*, 200–216. [[CrossRef](#)]
20. Rabuka, D.; Rush, J.S.; deHart, G.W.; Wu, P.; Bertozzi, C.R. Site-Specific Chemical Protein Conjugation Using Genetically Encoded Aldehyde Tags. *Nat. Protoc.* **2012**, *7*, 1052–1067. [[CrossRef](#)]
21. Gasteiger, E. ExpASY: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [[CrossRef](#)]
22. Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]
23. Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372*, 774–797. [[CrossRef](#)]
24. Schmidt, B.; Selmer, T.; Ingendoh, A.; von Figurat, K. A Novel Amino Acid Modification in Sulfatases That Is Defective in Multiple Sulfatase Deficiency. *Cell* **1995**, *82*, 271–278. [[CrossRef](#)]
25. Zheng, H.; Cooper, D.R.; Porebski, P.J.; Shabalin, I.G.; Handing, K.B.; Minor, W. CheckMyMetal: A Macromolecular Metal-Binding Validation Tool. *Acta Crystallogr. Sect. Struct. Biol.* **2017**, *73*, 223–233. [[CrossRef](#)]
26. Ponstingl, H.; Henrick, K.; Thornton, J.M. Discriminating between Homodimeric and Monomeric Proteins in the Crystalline State. *Proteins Struct. Funct. Bioinform.* **2000**, *41*, 47–57. [[CrossRef](#)]
27. van Loo, B.; Heberlein, M.; Mair, P.; Zinchenko, A.; Schüürmann, J.; Eenink, B.D.G.; Holstein, J.M.; Dilkaute, C.; Jose, J.; Hollfelder, F.; et al. High-Throughput, Lysis-Free Screening for Sulfatase Activity Using *Escherichia coli* Autodisplay in Microdroplets. *ACS Synth. Biol.* **2019**, *8*, 2690–2700. [[CrossRef](#)]
28. Pabis, A.; Duarte, F.; Kamerlin, S.C.L. Promiscuity in the Enzymatic Catalysis of Phosphate and Sulfate Transfer. *Biochemistry* **2016**, *55*, 3061–3081. [[CrossRef](#)] [[PubMed](#)]
29. Wu, J.; Filutowicz, M. Hexahistidine (His6)-Tag Dependent Protein Dimerization: A Cautionary Tale. *Acta Biochim. Pol.* **1999**, *46*, 591–599. [[CrossRef](#)]
30. Booth, W.T.; Schlachter, C.R.; Pote, S.; Ussin, N.; Mank, N.J.; Klapper, V.; Offermann, L.R.; Tang, C.; Hurlburt, B.K.; Chruszcz, M. Impact of an N-Terminal Polyhistidine Tag on Protein Thermal Stability. *ACS Omega* **2018**, *3*, 760–768. [[CrossRef](#)]
31. PyMOL. *The PyMOL Molecular Graphics System*, Version 2.0; Schrödinger, LLC: New York, NY, USA, 2016.
32. Krissinel, E.; Henrick, K. Secondary-Structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2256–2268. [[CrossRef](#)] [[PubMed](#)]
33. Krissinel, E.; Henrick, K. Multiple Alignment of Protein Structures in Three Dimensions. In *Computational Life Sciences*; Berthold, M.R., Glen, R.C., Diederichs, K., Kohlbacher, O., Fischer, I., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3695, pp. 67–78, ISBN 978-3-540-29104-6.
34. Holm, L. Using Dali for Protein Structure Comparison. In *Structural Bioinformatics*; Gáspári, Z., Ed.; Methods in Molecular Biology; Springer: New York, NY, USA, 2020; Volume 2112, pp. 29–42, ISBN 978-1-07-160269-0.
35. Laskowski, R.A.; Jabłońska, J.; Pravda, L.; Vařeková, R.S.; Thornton, J.M. PDBsum: Structural Summaries of PDB Entries. *Protein Sci.* **2018**, *27*, 129–134. [[CrossRef](#)]
36. Sillitoe, I.; Bordin, N.; Dawson, N.; Waman, V.P.; Ashford, P.; Scholes, H.M.; Pang, C.S.M.; Woodridge, L.; Rauer, C.; Sen, N.; et al. CATH: Increased Structural Coverage of Functional Space. *Nucleic Acids Res.* **2021**, *49*, D266–D273. [[CrossRef](#)] [[PubMed](#)]
37. Lewis, T.E.; Sillitoe, I.; Dawson, N.; Lam, S.D.; Clarke, T.; Lee, D.; Orengo, C.; Lees, J. Gene3D: Extensive Prediction of Globular Domains in Proteins. *Nucleic Acids Res.* **2018**, *46*, D435–D439. [[CrossRef](#)] [[PubMed](#)]
38. Golovin, A.; Henrick, K. MSDmotif: Exploring Protein Sites and Motifs. *BMC Bioinform.* **2008**, *9*, 312. [[CrossRef](#)] [[PubMed](#)]
39. Thompson, J.D.; Higgins, D.G.; Gibson, T.J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **1994**, *22*, 4673–4680. [[CrossRef](#)]
40. Sogi, K.M.; Gartner, Z.J.; Breidenbach, M.A.; Appel, M.J.; Schelle, M.W.; Bertozzi, C.R. Mycobacterium Tuberculosis Rv3406 Is a Type II Alkyl Sulfatase Capable of Sulfate Scavenging. *PLoS ONE* **2013**, *8*, e65080. [[CrossRef](#)]
41. Jorge, A.M.; Schneider, J.; Unsleber, S.; Xia, G.; Mayer, C.; Peschel, A. Staphylococcus Aureus Counters Phosphate Limitation by Scavenging Wall Teichoic Acids from Other Staphylococci via the Teichoicase GlpQ. *J. Biol. Chem.* **2018**, *293*, 14916–14924. [[CrossRef](#)]
42. Holder, P.G.; Jones, L.C.; Drake, P.M.; Barfield, R.M.; Bañas, S.; de Hart, G.W.; Baker, J.; Rabuka, D. Reconstitution of Formylglycine-Generating Enzyme with Copper(II) for Aldehyde Tag Conversion. *J. Biol. Chem.* **2015**, *290*, 15730–15745. [[CrossRef](#)]
43. Woo, L.W.L.; Ganeshapillai, D.; Thomas, M.P.; Sutcliffe, O.B.; Malini, B.; Mahon, M.F.; Purohit, A.; Potter, B.V.L. Structure-Activity Relationship for the First-in-Class Clinical Steroid Sulfatase Inhibitor Irosustat (STX64, BN83495). *ChemMedChem* **2011**, *6*, 2019–2034. [[CrossRef](#)]

44. Pérez-Jiménez, M.M.; Monje-Moreno, J.M.; Brokate-Llanos, A.M.; Venegas-Calderón, M.; Sánchez-García, A.; Sansigre, P.; Valladares, A.; Esteban-García, S.; Suárez-Pereira, I.; Vitorica, J.; et al. Steroid Hormones Sulfatase Inactivation Extends Lifespan and Ameliorates Age-Related Diseases. *Nat. Commun.* **2021**, *12*, 49. [[CrossRef](#)]
45. Otwinowski, Z.; Minor, W. Processing of X-Ray Diffraction Data Collected in Oscillation Mode. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 1997; Volume 276, pp. 307–326, ISBN 978-0-12-182177-7.
46. Vagin, A.; Teplyakov, A. MOLREP: An Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025. [[CrossRef](#)]
47. Minor, W.; Cymborowski, M.; Otwinowski, Z.; Chruszcz, M. HKL-3000: The Integration of Data Reduction and Structure Solution—from Diffraction Images to an Initial Model in Minutes. *Acta Crystallogr. D Biol. Crystallogr.* **2006**, *62*, 859–866. [[CrossRef](#)]
48. Bienert, S.; Waterhouse, A.; de Beer, T.A.P.; Tauriello, G.; Studer, G.; Bordoli, L.; Schwede, T. The SWISS-MODEL Repository—New Features and Functionality. *Nucleic Acids Res.* **2017**, *45*, D313–D319. [[CrossRef](#)] [[PubMed](#)]
49. Guex, N.; Peitsch, M.C.; Schwede, T. Automated Comparative Protein Structure Modeling with SWISS-MODEL and Swiss-PdbViewer: A Historical Perspective. *Electrophoresis* **2009**, *30*, S162–S173. [[CrossRef](#)] [[PubMed](#)]
50. Waterhouse, A.; Bertoni, M.; Bienert, S.; Studer, G.; Tauriello, G.; Gumienny, R.; Heer, F.T.; de Beer, T.A.P.; Rempfer, C.; Bordoli, L.; et al. SWISS-MODEL: Homology Modelling of Protein Structures and Complexes. *Nucleic Acids Res.* **2018**, *46*, W296–W303. [[CrossRef](#)]
51. Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC 5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [[CrossRef](#)]
52. Painter, J.; Merritt, E.A. TLSMD Web Server for the Generation of Multi-Group TLS Models. *J. Appl. Crystallogr.* **2006**, *39*, 109–111. [[CrossRef](#)]
53. Chen, V.B.; Arendall, W.B.; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12–21. [[CrossRef](#)]
54. Handing, K.B.; Niedzialkowska, E.; Shabalin, I.G.; Kuhn, M.L.; Zheng, H.; Minor, W. Characterizing Metal-Binding Sites in Proteins with X-Ray Crystallography. *Nat. Protoc.* **2018**, *13*, 1062–1090. [[CrossRef](#)]
55. Zheng, H.; Chordia, M.D.; Cooper, D.R.; Chruszcz, M.; Müller, P.; Sheldrick, G.M.; Minor, W. Validation of Metal-Binding Sites in Macromolecular Structures with the CheckMyMetal Web Server. *Nat. Protoc.* **2014**, *9*, 156–170. [[CrossRef](#)]
56. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 486–501. [[CrossRef](#)] [[PubMed](#)]
57. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)] [[PubMed](#)]
58. Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L.E.; Brookes, D.H.; Wilson, L.; Chen, J.; Liles, K.; et al. Improvements to the APBS Biomolecular Solvation Software Suite. *Protein Sci.* **2018**, *27*, 112–128. [[CrossRef](#)] [[PubMed](#)]
59. 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. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)] [[PubMed](#)]
60. Ellis, K.J.; Morrison, J.F. [23] Buffers of Constant Ionic Strength for Studying PH-Dependent Processes. In *Methods in Enzymology*; Elsevier: Amsterdam, The Netherlands, 1982; Volume 87, pp. 405–426, ISBN 978-0-12-181987-3.