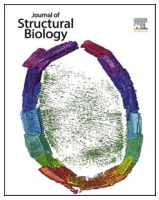
Journal of Structural Biology 217 (2025) 108260

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi



Research Article

Crystal structure of *Escherichia coli* type I signal peptidase P84A in complex with lipopeptide antibiotic arylomycin A_2

Chuanyun Luo, Mark Paetzel *

Department of Molecular Biology and Biochemistry, Simon Fraser University, South Science Building 8888 University Drive, Burnaby, British Columbia V5A 1S6 Canada

ARTICLE INFO

Editor: Bauke Dijkstra

Keywords:
Signal peptidase
Signal peptide
Arylomycin
Lipopeptide inhibitor
Antibiotic
Protease
Preprotein
Preprotein binding

ABSTRACT

Type I signal peptidase (SPase I) is an essential membrane-bound enzyme that removes amino-terminal signal peptides from secretory proteins. Owing to its critical role in bacterial viability and its periplasmic accessibility, SPase I has emerged as an attractive target for antibiotic development. Arylomycins, a class of macrocyclic lipohexapeptide natural products, inhibit SPase I by binding to its active site. Previous studies have identified a key resistance determinant—a proline residue at the base of the substrate-binding groove (Pro84 in *Escherichia coli* SPase I)—which reduces arylomycin affinity. Here, we present the crystal structure of the *E. coli* SPase I P84A mutant in complex with arylomycin A_2, revealing that the introduced alanine enables an additional hydrogen bond between the enzyme backbone and the arylomycin N-terminal carbonyl, thus enhancing the affinity for arylomycins. Furthermore, a newly developed preprotein-binding assay utilizing a non-cleavable version of ProOmpA Nuclease A demonstrates that substituting SPase I Pro84 with serine or leucine disrupts substrate recognition, underscoring the delicate balance between inhibitor resistance and substrate processing. These findings reveal that residue Pro84 participates in the interaction between preprotein signal peptides and the *E. coli* SPase I substrate-binding groove, offering a foundation for designing next-generation arylomycin analogs with improved antibacterial potency.

1. Introduction

1.1. Signal peptidase

Most secretory proteins are synthesized with an amino-terminal signal peptide that functions in targeting to and translocation across the cellular membrane via a secretion translocase apparatus. In bacteria this process occurs predominately post-translationally (Kaushik et al., 2022).

Type 1 signal peptidase (SPase I) is responsible for cleaving off the signal peptide, allowing the mature secretory protein to leave the membrane and perform its intended function. SPase I has been shown to be essential and is being investigated as a potential antibiotic target (Date, 1983). The most thoroughly investigated SPase I is that of *Escherichia coli*. This monomeric membrane-bound enzyme has two amino-terminal transmembrane segments, a cytoplasmic loop and a carboxy-terminal catalytic domain that resides in the periplasmic space. The catalytic mechanism of bacterial SPase I utilizes a serine nucleophile and a lysine general-base that has been investigated by site-directed mutagenesis (Sung and Dalbey, 1992; Tschantz et al., 1993), chemical modifications (Paetzel et al., 1997) and crystallographic analysis (Paetzel et al., 1998). Further information on *E. coli* SPase I and its substrates and inhibitors can be found in a review (Paetzel, 2014). The *E. coli* SPase I construct used in this study is the periplasmic catalytic domain of the enzyme, referred to as Δ2-76, that lacks the two amino-terminal transmembrane segments and the small cytoplasmic loop that links the transmembrane segments (Kuo et al., 1993; Paetzel et al., 1995; Tschantz et al., 1995). The sequence for *E. coli* SPase I Δ2-76 is shown in the supplementary material.

1.2. Arylomycins

Arylomycins are natural products first isolated from *Streptomyces* extracts (Holtzel et al., 2002; Schimana et al., 2002). They are secondary metabolites produced by non-ribosomal peptide synthesis and have been shown to have antibiotic activity based on their ability to inhibit SPase I (Kulanthaivel et al., 2004). The lipohexapeptide arylomycin A_2 has the sequence: N-methyl D-serine, D-alanine, Glycine, N-methyl-4-

* Corresponding author.
E-mail address: mpaetzel@sfu.ca (M. Paetzel).

https://doi.org/10.1016/j.jsb.2025.108260
Received 30 June 2025; Received in revised form 5 November 2025; Accepted 6 November 2025
Available online 7 November 2025

hydroxyphenylglycine, L-alanine, L-tyrosine, with a 12-carbon atom branched fatty acid (isoC12) attached via an amide bond to the amino terminus. The N-methyl-4-hydroxyphenylglycine residue is cross-linked via the *ortho*-carbon atom of its phenol ring to the *ortho*-carbon atom in the phenol ring of the tyrosine residue forming a (3,3)-biaryl bridge and resulting in a three-residue macrocycle (Fig. 1). New synthetic routes have allowed for modifications that facilitate optimization of the arylomycins (Peters et al., 2018; Smith et al., 2018).

1.3. Arylomycin susceptible species

Arylomycins have activity against *Streptococcus pneumonia* (Kulanthaivel et al., 2004), some soil bacteria (Schimana et al., 2002), and *Staphylococcus epidermidis* (Roberts et al., 2007) but most bacterial species tested are resistant to the lipopeptide (Smith et al., 2010). Smith et al. combined genetics, biochemical and phylogenetic analysis to show that a proline residue located at the base of the substrate binding surface of SPase I is responsible for arylomycin resistance. This residue in *E. coli* SPase I is Pro84, and proline at this position (*E. coli* SPase I numbering) is frequently found among *E. coli* SPase I homologues in both Gram-negative and Gram-positive bacteria. The higher the similarity to *E. coli* SPase I (i.e., percentage identity), the more frequently proline is found at this position (Fig. 2).

Smith et al. showed that a selection experiment can be used to discover *S. epidermidis* mutants that can grow in the presence of arylomycin. They showed that a serine to proline mutation at position 29 in signal peptidase I (which corresponds to Pro84 in *E. coli* SPase I) was responsible for evolving resistance to arylomycin. They then developed mutant strains of naturally arylomycin resistant bacteria such as *E. coli*, *Pseudomonas aeruginosa*, and *Staphylococcus aureus* (contain proline at the 84 position, *E. coli* numbering) and converted them to arylomycin-sensitive strains by mutating residue 84 to a non-proline residue. They replaced the residue at the 84 position with the 19 other natural amino acid residues and showed that only proline results in arylomycion resistance. Their SPase I / arylomycin binding assays showed that *E. coli* SPase I P84S has a higher affinity to arylomycin in comparison to *E. coli* SPase I with Pro84. They observed the same result using the *S. aureus* enzyme. Phylogenetic analysis of all five phyla of sequenced bacteria along with sampling of arylomycin sensitivity was performed. Interestingly, they observed that *Helicobacter pylori* has an alanine at the 84 position, *E. coli* numbering, and is arylomycin sensitive.

Previous crystal structures of *E. coli* SPase I in complex with arylomycin A_2 show that *E. coli* SPase Pro84 is located adjacent to the arylomycin amide carbonyl oxygen that links the lipid to the amino-terminus of the peptide (Liu et al., 2011; Luo et al., 2009; Paetzel et al., 2004). Proline residues within a polypeptide chain do not have a mainchain amide hydrogen to function as a hydrogen bond donor.

1.4. Preprotein substate PONA and its non-cleavable version PONA A22P

The preprotein substrate Pro-OmpA nuclease A (PONA) is the most characterized and used preprotein substrate for bacterial type I signal peptidases. It is a chimera of the signal peptide from *E. coli* outer membrane protein A (OmpA, residues 1–21) and the mature region of *Staphylococcus aureus* nuclease A. PONA was developed by Takahara et al. (Takahara et al., 1985) and used by Chatterjee et al. (Chatterjee et al., 1995) to determine Michaelis-Menten parameter for *E. coli* SPase I. The k_{cat}/K_M for PONA was shown to be six-orders of magnitude greater than peptide substrate used at the time. Although this is a sensitive and the most specific substrate for measuring SPase I activity, preproteins are difficult to express and purify and the SDS-PAGE band-shift assay is difficult and laborious to perform. The sequence for PONA is shown in the supplementary material.

A proline at the P1′ (+1) position within a preprotein relative to the SPase I cleavage site prevents cleavage and produces a SPase I competitive inhibitor. This has been observed using pre-beta-lactamase (Pluckthun and Knowles, 1987), pre-maltose-binding protein (Barkocy-Gallagher and Bassford, 1992), and in a *sec*-independent preprotein (Nilsson and von Heijne, 1992). We have applied the + 1 Pro strategy using PONA to create the competitive inhibitor PONA-A22P. We have developed a novel SPase I / preprotein binding assay using PONA-A22P with Surface Plasmon Resonance (SPR).

We have determined the crystal structure of a P84A mutant of *E. coli* SPase I Δ2-76 in complex with arylomycin A_2. Electron density within the substrate binding groove that interacts with the lipopeptide inhibitor shows the additional hydrogen bond predicted to play a role in arylomycin A_2 susceptibility. Using a preprotein binding assay, we found that substituting SPase I Pro84 with serine or leucine reduces the binding affinity for the non-cleavable preprotein proOmpA Nuclease A (PONA-A22P). These results indicate that Pro84 contributes to the interaction surface between the *E. coli* SPase I binding groove and the signal peptide of preprotein substrates.

2. Materials and methods

2.1. DNA construct preparation

E. coli SPase Δ2-76 site-directed mutants (P84A, P84L, and P84S) used in this study were constructed by a standard PCR method, using the construct pET3d-*E. coli* SPase Δ2-76 as the primary template (Tschantz

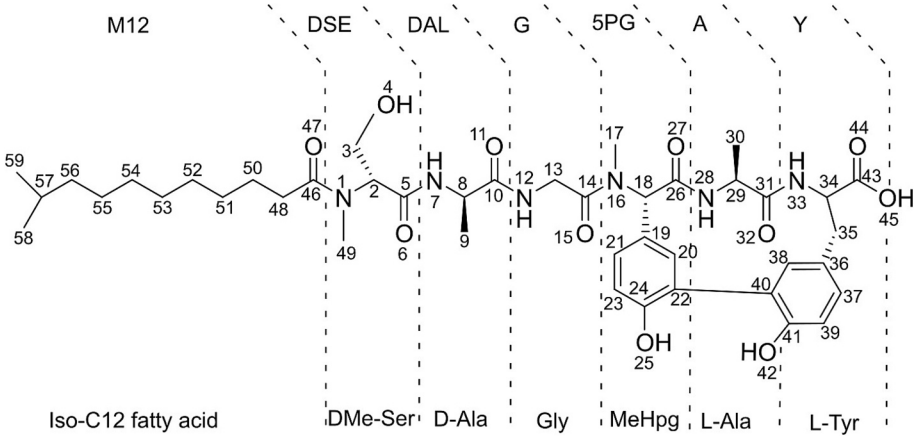


Fig. 1. SPase I inhibitor and lipohexapeptide antibiotic arylomycin A_2. All atoms are labeled. The top labels correspond to the nomenclature for each part of arylomycin A_2 (PRD_000117) in the PDB database. The bottom labels correspond to the nomenclature for each part of arylomycin A_2 as reported in previous literature.

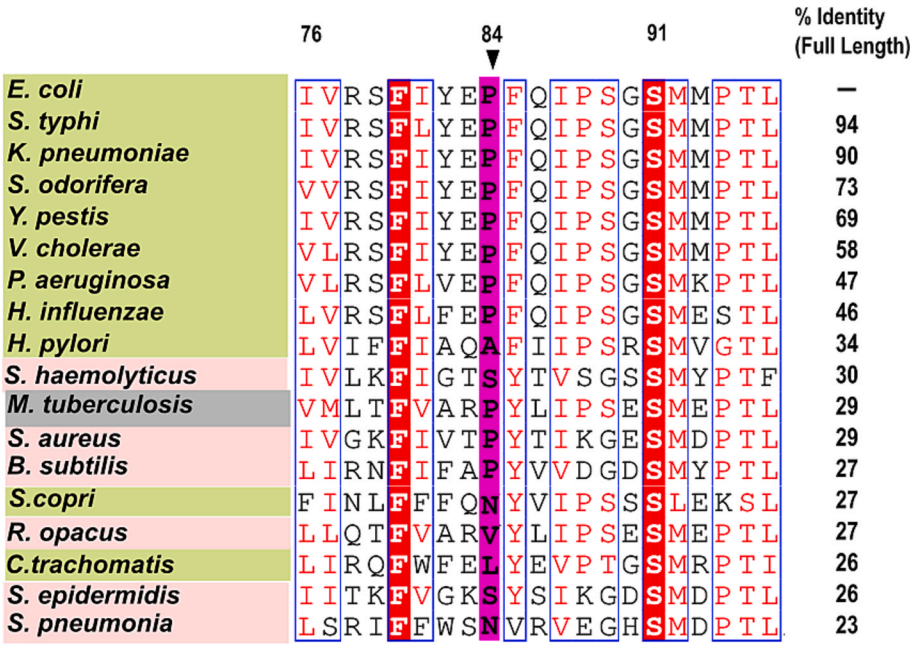


Fig. 2. Sequence alignment displaying the residue diversity at Pro84 *(E. coli* SPase I numbering) from homologues in Gram-negative and Gram-positive bacteria. Conserved residues are shown with red background, similar residues are shown in red text with blue boxes. The sequences are ordered based on their level of identity to the *E. coli* SPase I sequence. Gram-negative bacteria are highlighted in yellow while gram-positive bacteria are highlighted in pink. Bacteria that are classified as acid-fast Gram-positive bacteria are highlighted in grey. The protein sequences were acquired from protein knowledgebase (UniProtKB): *Escherichia coli* (P00803); *Salmonella typhi* (P0A1W3); *Klebsiella pneumoniae* (B5XNG9); *Serratia odorifera* (D4E2Q4); *Yersinia pestis* (A0A380PM08); *Vibrio cholerae* (Q9KPB1); *Pseudomonas aeruginosa* (Q9I5G7); *Haemophilus influenzae* (P44454); *Helicobacter pylori* (O25300); *Staphylococcus haemolyticus* (Q4L306); *Mycobacterium tuberculosis* (P9WKA0); *Staphylococcus aureus* (Q5HHB9); *Bacillus subtilis*(P28628); *Segatella copri* (D1PGA3); *Rhodococcus opacus* (C1B2S5); *Chlamydia trachomatis* (O84023); *Staphylococcus epidermidis* (Q5HQJ6); *Streptococcus pneumoniae* (O07344).

et al., 1995). The *E. coli* SPase Δ2-76 construct used in this study lacks the two amino-terminal transmembrane segments and represents the periplasmic catalytically active domain. This construct includes the amino-terminal methionine followed by residues 77 to 324 (UniProt accession #: P00803). The synthetic primers used include: P84A forward primer 5′-GTTCGTTTATTTATGAAGCGTTCCAGATCCCGTC-3′ and P84A reverse primer 5′-GACGGGATCTGGAACGCTTCATAAA-TAAACGAAC-3′; P84L forward primer 5′-GTTCGTTTATTTAT-GAACTGTTCCAGATCCCGTCAG-3′ and P84L reverse primer 5′-CTGACGGGATCTGGAACAGTTCATAAATAAACGAAC-3′; P84S forward primer 5′-GTGCGTTCGTTTATTTATGAAAGCTTCCAGATCCCGTCAGGT TCG-3′ and P84S reverse primer 5′-CGAACCTGACGGGATCTG-GAAGCTTTCATAAATAAACGAACGCAC-3′. The resultant mutants were confirmed by DNA sequencing.

The preprotein substrate pro-OmpA NucleaseA (PONA) gene was amplified via a standard PCR using pET21a-PONA (a gift from the Dr. Ross Dalbey laboratory, Ohio State University) as a template (Carlos et al., 2000). The oligonucleotides used for amplifying the PONA gene were forward primer 5́ GGC ATA TGA AAA AGA CAG CTA TCG CGA TTG CAG 3′ and revers primer 5́ GGC TCG AGT TAT TGA CCT GAA TCA GCG TTG TCT TC 3′ and then was cloned into the expression vector pET28a (Novagen) using the NdeI and XhoI restriction sites to create pET28a-PONA. The pET28a-PONA construct was verified and confirmed by DNA sequencing. The expressed protein contains an amino-terminal *hexa*-histidine nickel affinity tag. The sequence for the expressed protein is provided in the supplementary material.

PONA was converted into a non-cleavable competitive inhibitor by mutating the P1′ position alanine to a proline (PONA-A22P) via site-directed mutagenesis using a standard PCR method. The construct pET28a-PONA was used as a template. The synthetic primers used included forward primer 5′ −GCTACCGTAGCGCAGGCCCCGACTTCAACTAAAAAATTAC-3′ and reverseprimer 5′ −GTAATTTTTTAGTTGAAGTCGGGGCCTGCGCTA CGGTAGC-3′. The mutation A22P was confirmed by DNA sequencing.

2.2. Protein expression and purification

The plasmids coding for SPase Δ2-76 and the mutant SPase Δ2-76 P84A, SPase Δ2-76 P84L, and SPase Δ2-76 P84S were transformed into *E. coli* BL21(λDE3) cells for protein expression. *E. coli* SPase Δ2-76 and the mutants were expressed and purified as previously described (Paetzel et al., 1995). Briefly, the culture was grown in LB medium with ampicillin (100 μg/ml) at 37 °C to an OD_{600} of 0.6, induced with isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The induced cells were lysed using three pulses of sonication, each pulse lasting 15 s. with an amplitude of 30 % and separated by a 30 s. rest interval (Fisher Scientific sonic dismembrator Model 500). The sonicated cells were homogenized using an Avestin Emulsiflex-3C homogenizer at 15,000–20,000 psi for 5 min.

The overexpressed protein accumulated as inclusion bodies, which were isolated by centrifugation and then dissolved in denaturing buffer (4 M guanidine hydrochloride in 20 mM Tris-HCl, pH 7.4 buffer) and refolded in refolding buffer (20 mM Tris-HCl, pH 7.4, 10 mM EDTA, and 0.5 % Triton X-100), and the soluble protein was dialyzed against the dialysis buffer (20 mM Tris-HCl, pH 7.4, and 0.1 % Triton X-100).

The protein was purified using self-packed anion-exchange column with Q Sepharose fast Flow resin (Amersham Biosciences) and eluted by applying a high concentration salt solution of 0.7 M NaCl in 20 mM Tris-HCl, pH 7.4, and 0.1 % Triton X-100, and further purified using gel filtration chromatography with a Sephadex 200 column (GE Healthcare), equilibrated in buffer 20 mM Tris-HCl pH 7.4, 100 mM NaCl, and 0.1 % Triton X-100. The pure protein fractions were pooled and concentrated using an ultrafiltration concentrator (Amicon) through an YM-10 membrane (Amicon, MWCO 10KDa, Millipore Sigma). The protein concentration was measured by absorbance at 280 nm using a Nanodrop ND-100 spectrophotometer. The extinction coefficient

(38515 M^{-1} cm^{-1}) for SPase Δ2-76 was calculated based on the amino acid sequence using ProtParam (Gasteiger et al., 2005).

The expression plasmids coding for PONA and PONA-A22P were transformed into *E. coli* BL21(λDE3) cells. The cell growth, induction and cell-lysis steps were performed using the same protocol as described above. Both expressed proteins have an amino-terminal *hexa*-histidine affinity tags and were purified using Ni^{+2} affinity chromatography. In brief, the cell lysate was clarified by centrifugation (35,000xg) for 40 min at 4 °C. The supernatant was loaded on to a nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen) column pre-equilibrated with buffer A (20 mM Tris-HCl, pH 8.0; 100 mM NaCl; 0.01 % n-dodecyl β-D-maltoside DDM) and the protein was eluted from the column using a step gradient (100–500 mM imidazole in buffer 20 mM Tris-HCl, pH 8.0; 100 mM NaCl; 0.01 % n-dodecyl β-D-maltoside DDM in 100 mM increments). The fractions (200–400 mM) containing pure protein were pooled and dialyzed against buffer A overnight at 4 °C to remove imidazole. Dialyzed protein was then concentrated using an ultrafiltration concentrator (Amicon) through an YM-10 membrane (Amicon, MWCO 10KDa, Millipore Sigma). Further purification was performed using size-exclusion chromatography with a Sephadex 200 column (GE Healthcare), equilibrated in buffer A attached to an AKTA FPLC flowing at 1 mL/min. The pure protein fractions were pooled and concentrated again using the above ultrafiltration concentrator. The protein concentration was measured by absorbance at 280 nm using a Nanodrop ND-100 spectrophotometer. The extinction coefficient (15,930 M^{-1} cm^{-1} for both PONA and PONA-A22P) was calculated based on the PONA and PONA-A22P amino acid sequence using ProtParam (Gasteiger et al., 2005).

2.3. Co-crystallization and data collection

Prior to co-crystallization, SPase Δ2-76 P84A (31.2 mg/ml in 20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 0.1 % Triton-X100) was combined with the arylomycin A_2 (10 mM in DMSO) in a 1:1 M ratio and incubated on ice for one hour.

Co-crystallization trials for SPase Δ2-76 P84A in complex with arylomycin A_2 were carried out using the sitting-drop vapor diffusion method. The final optimized reservoir condition that produced diffraction quality crystals for data collection was: 25 % v/v PEG 4000, 0.05 M ammonium acetate, 0.1 M sodium acetate pH 4.6. The drop was made up of 2 μl of the protein/inhibitor complex described above, 2 μl of reservoir solution, and 2 μl of 0.1 M L-proline. This 6 μl drop was equilibrated over 1 ml of reservoir solution at 21°C. The crystals formed from a light precipitate after approximately two weeks and had an average size of ~ 0.1 × 0.5 × 0.05 mm.

Before data collection, the crystal was transferred by a pipet from the growth drop to a cryoprotectant composed of 25 % w/v PEG 4000, 0.05 M ammonium acetate, 0.1 M sodium acetate pH 4.6, 20 % v/v 1,2-ethanediol) for 30 s. The crystal was mounted on a Hampton Research loop and cryo-cooled by being directly placed into liquid nitrogen and stored in a dewer until data collection. The diffraction data set was collected at Beamline 081D-1 of the Canadian Light Source (CLS), Canadian Macromolecular Crystallography Facility using a wavelength of 1.1047Å. The crystal-to-detector distance was set to 250 mm. All frames (457) were recorded on a MARMOSAIC Rayonix MX300 CCD detector with a 0.5° oscillation angle and an exposure time of 1 s. per frame. The data revealed diffraction beyond a resolution of 2.32 Å. Data were indexed and scaled using the programs MOSFLM (Battye et al., 2011) and SCALA (Evans, 2006). The crystals belong to the tetragonal space group $P4_32_12$ and the unit cell dimensions were determined to be a = 71.80 Å, b = 71.80 Å, and c = 263.31 Å. The Matthews coefficient (V_m) is 3.03 $Å^3$/Da for two molecules in the asymmetric unit. The fraction of the crystal volume occupied by solvent was 59.5 %, as calculated by the program *Matthews_coef* in the CCP4i suite of programs (Kantardjieff and Rupp, 2003; Matthews, 1968). For crystal and data collection statistics see Table 1.

2.4. Structure determination

A molecular replacement solution was found using the program Phaser (McCoy, 2007). The atomic coordinates used for the search model were taken from a 2.4 Å crystal structure of SPase Δ2-76, protein chain B within PDB code 3S04 (Liu et al., 2011). The topology and parameter files for the inhibitor were generated using the program eLBOW (Moriarty et al., 2009). Coordinates for the inhibitor were manually docked into clear electron difference density ($F_o - F_c$) near the active site and substrate binding groove. Residue Pro 84 was mutated to Ala 84 to fit the electron density.

Model building and analysis was performed using the program Coot (Emsley and Cowtan, 2004). Refinement of the structure was carried out using the programs CCP4 Refmac (Kovalevskiy et al., 2018; Murshudov et al., 1999; Vagin et al., 2004; Winn et al., 2003) and PHENIX (Adams et al., 2002; Adams et al., 2004; Adams et al., 2010; Murshudov et al., 2011; Vagin et al., 2004). Throughout refinement 3.6 % of the reflections were set aside for cross-validation. Water molecules were added to well-defined peaks (2.0 σ and greater in the $F_o - F_c$ maps). Final refinement and analysis statistics of the complex are provided in Table 1. The stereochemistry of the structure model was analyzed with the program PROCHECK (Laskowski et al., 1993). The atomic coordinates (accession code: 9NLO) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ.

Table 1
Crystallographic statistics.

PDB accession code	9NLO
Crystal Parameters	
Space group	$P4_32_12$
a,b,c (Å)	71.80, 71.80, 263.31
Solvent content (%)	59.5
Matthew's coefficient ($Å^3$/Da)	3.03
Data Collection	
Wavelength (Å)	1.1047
Resolution (Å)	52.63 – 2.32(2.45 – 2.32)[a]
Total reflections	221,402 (32122)
Unique reflections	30,863 (4381)
R_{merge} [b]	0.097 (0.392)
R_{pim}	0.039 (0.154)
Mean (*I*)/σ (*I*)	13.0 (5.8)
Completeness (%)	99.7 (99.8)
Redundancy	7.2 (7.3)
Refinement	
Protein chains in A.U.	2
Residues	455
Inhibitors	2
Ligands	2
Water molecules	124
Total number of atoms	3726
R_{cryst} [c] / R_{free} [d] (%)	20.9/23.4
Average B-factor ($Å^2$) (all atoms)	46.0
Average B-factor ($Å^2$) (protein)	46.8
Average B-factor ($Å^2$) (solvent)	42.3
Average B-factor ($Å^2$) (inhibitor)	49.7
Rmsd on angles (°)	1.400
Rmsd on bonds (Å)	0.006
Ramachandran	
favored (%)	95.24
allowed (%)	4.76
outliers (%)	0

[a] Values in brackets represents the highest resolution shell data.

[b] $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an individual reflection and $\langle I(hkl)\rangle$ is the mean intensity of that reflection.

[c] $R_{cryst} = \sum_{hkl}||F_{obs}| - |F_{calc}|| / \sum_{hkl}|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively.

[d] R_{free} is calculated using 3.6 % of the reflections randomly excluded from refinement.

2.5. Preprotein binding assay

The surface plasmon resonance (SPR) experiments were performed on *E. coli* SPase Δ2-76 and *E. coli* SPase Δ2-76 mutants binding on PONA-A22P using a BIACORE X100 (GE Healthcare) at 25℃. A commercially available NTA sensor chip (GE Healthcare) was used. The proteins, protein substrate inhibitor PONA-A22P, and solutions were prepared using the Biacore running buffer B (10 mM HEPES; 150 mM NaCl; 0.005 % v/v Tween 20; 50 μM EDTA; pH 7.4). The NTA sensor chip was primed, conditioned, and immobilized using the manufacturer's protocol for NTA chips. In brief, the NTA sensor chip was conditioned with 0.35 M EDTA. In each cycle, 0.5 mM $NiCl_2$ was injected at 10 μL/min for 3 min, followed by injection of PONA-A22P (50 nM) at 5 μl/min for 150 s. The unattached PONA-A22P was washed with 3 mM EDTA. The capture level of PONA-A22P (50 nM) was around 100 RU (resonance units). Then, pure *E. coli* SPase Δ2-76 or *E. coli* SPase Δ2-76 mutant proteins (>95 % based on SDS-PAGE gel results) were allowed to associate separately with PONA-A22P for 150 s and dissociated for 300 s at the flow rate of 10 μl/min. Again, after every measurement at every concentration, the entire immobilized PONA-A22P surface was removed using 0.35 M EDTA at 10 μl/min. The regeneration solution (0.35 M EDTA) was passed over the immobilized PONA-A22P surface for a contact time of 60 s and the sensor chip was then recoated with a fresh PONA-A22P solution for the next measurement. The analysis was performed three time or more and the error values calculated as summarized in table 2.

2.6. Figure preparation

Figures were prepared using the programs ChemDraw (version 23.1.2 Revvity Signals Software Inc.), PyMol (DeLano, 2002) and ChimeraX (Goddard et al., 2018; Meng et al., 2023). The alignment figure was prepared using the programs ClustalW2 and ESPript 2.2 (Gouet et al., 2003; Larkin et al., 2007).

3. Results and discussion

3.1. X-ray crystal structure of E. coli SPase I Δ2-76 P84A in complex with arylomycin A_2

Mutant forms of *E. coli* SPase I catalytic domain (Δ2-76) with alanine, serine and leucine at residue position 84 were expressed and purified. All enzymes were subjected to co-crystallization screening with arylomycin A_2 but only the mutant enzyme P84A produced diffraction quality co-crystals with the inhibitor. Diffraction data were collected, processed and a structure was determined based on molecular

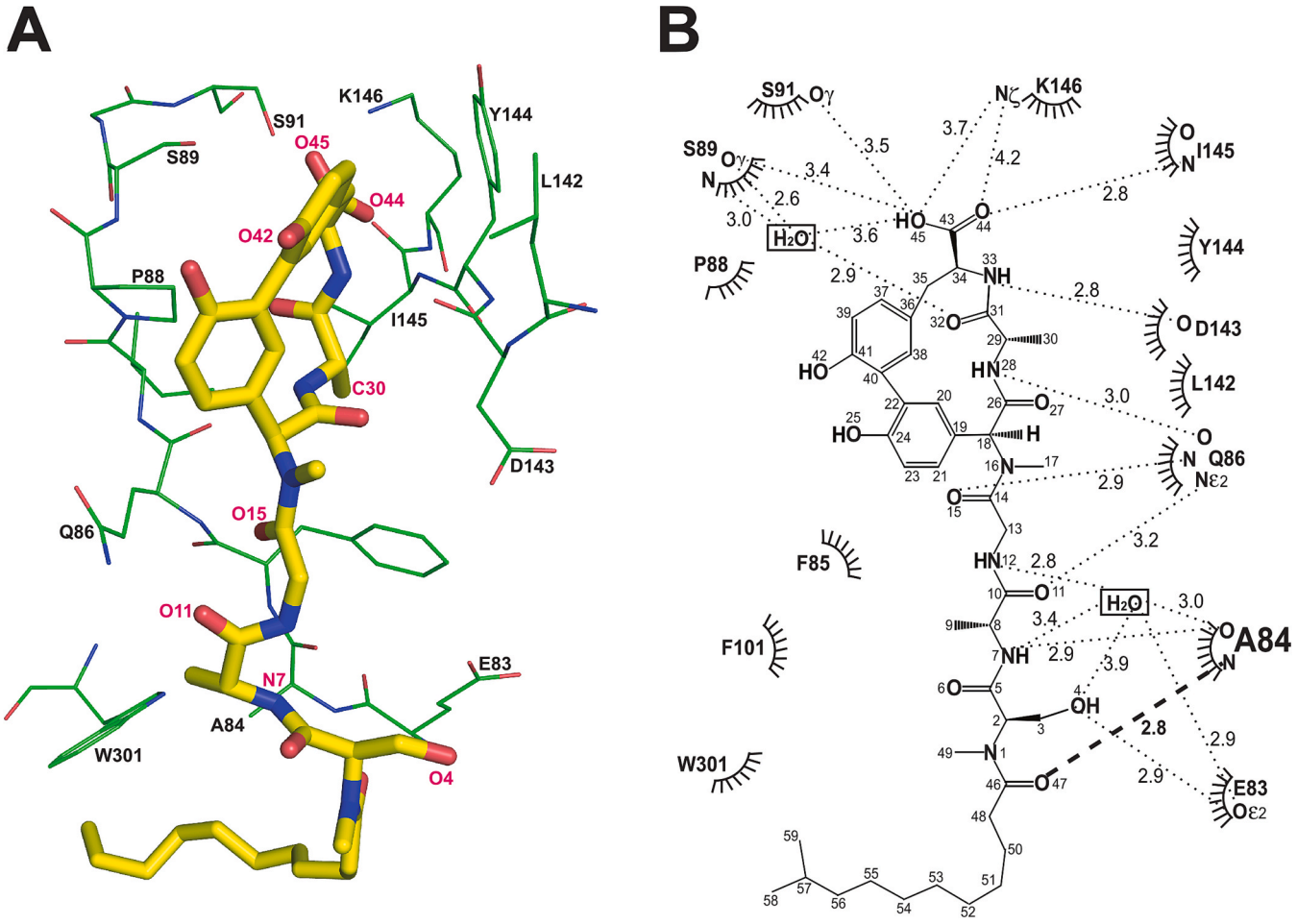


Fig. 3. View of arylomycin A_2 within the active site and substrate-binding groove of SPase. **A**. SPase is shown in line representation and colored by element (carbon, green; oxygen, red; nitrogen, blue; residues labeled in black). Arylomycin A_2 is shown in stick and colored by element (carbon, yellow; nitrogen, blue; oxygen, red; atoms labeled in red). Chain A within the asymmetric unit was used to make the figure. The partial (up to C11) fatty acid tail of arylomycin A_2 is included but the rest of it (C12) is not included due to weak electron density. **B**. Schematic representation of arylomycin A_2 interactions with the active site and substrate binding groove of SPase. Dashed lines indicate polar interactions and distances are in angstroms (average value for the two molecules in the asymmetric unit). An additional hydrogen bond (2.8Å) that is formed solely due to the introduced mutant P84A is highlighted with thick dashed line. The residues involved in van der Waals interactions are symbolized by arches.

replacement phasing. Clear difference density was observed for all arylomycin A_2 amino acid residues and nearly all the lipid atoms. The electron density for residue 84 was consistent with the mutation to alanine. The model for the lipopeptides was fit into the positive difference density and adjustments to the enzyme's polypeptide chain were made, and the structure was refined to 2.3 Å resolution (Fig. S1). Two regions of positive difference density were observed adjacent to the arylomycin, but they were too small and irregular to successfully refine a solvent molecule (Fig. S2). See Table 1 for crystallographic statistics.

The molecular environment and binding interactions between arylomycin A_2 and the substrate-binding groove residues of *E. coli* SPase I Δ2-76 P84A are shown in Fig. 3 and the polar interactions are summarized in Table S1. As observed previously, the C-terminal L-tyrosine carboxylate group of arylomycin A_2 makes direct polar contacts with all the catalytic residues in the active site of *E. coli* SPase I and the peptide's mainchain makes productive polar interactions in a parallel β-sheet fashion with the residues that make up the substrate binding groove (Liu et al., 2011; Luo et al., 2009; Paetzel et al., 2004). Also shown in Fig. 3B and listed in Table S2 are the polar interactions of two conserved water molecules within the substrate-binding groove of *E. coli* SPase I Δ2-76 P84A in complex with arylomycin A_2. Designing new functional groups within arylomycin that displace these waters may be a helpful strategy to improve the inhibitor.

Superposition of the active site residues of this SPase I P84A arylomycin A_2 complex with previous SPase I Pro84 – arylomycin A_2 complex structures reveal a very similar binding mode for the lipopeptide (Fig. S3)(Liu et al., 2011; Luo et al., 2009; Paetzel et al., 2004). A table of the arylomycin A_2 structure superposition root mean square deviation values from three different SPase I – arylomycin A_2 complex crystal structures (1t7d: *E. coli* SPase I Pro84 arylomycin A_2 complex (Paetzel et al., 2004), 3iiq: *E. coli* SPase I Pro84 arylomycin A_2 β-sultam ternary complex(Luo et al., 2009), and 9nlo: *E. coli* SPase I P84A arylomycin A_2 complex) shows that the lipopeptide maintains a consistence overall conformation within the enzymes substrate binding groove (Fig. S3, Tables S3 and S4).

A direct comparison between the mutant *E. coli* SPase I P84A arylomycin A_2 complex reported here and a previously determined *E. coli* SPase I Pro84 arylomycin A_2 complex reveals a hydrogen bond (2.8 Å) between mainchain NH of P84A to atom O47 of arylomycin A_2. The atom O47 is the carbonyl oxygen of the fatty acid that is attached to the amino-terminal N-methyl-D-serine (Fig. 4).

The distance between the mainchain nitrogen at position 84 and O47 of arylomycin A_2 in previously determined *E. coli* SPase I Pro84 structures (PDB 3iiq and PDB 1t7d) is significantly longer, on average 4.7 Å. The greater distance and lack of a hydrogen bonding donor NH group in Pro84 leads to the lack of contact at this position within the arylomycin A_2 SPase I interaction.

3.2. How do mutations at position 84 of *E. coli* SPase I affect preprotein substrate binding?

As discussed previously a proline at the + 1 or P1′ position relative to the signal peptide scissile bond is known to prevent cleavage by SPase I and produce a competitive SPase I inhibitor (Barkocy-Gallagher and Bassford, 1992; Nilsson and von Heijne, 1992). We have mutated Ala22 to proline in the preprotein substrate pro-OmpA Nuclease A (PONA) to produce the non-cleavable pre-protein competitive inhibitor PONA-A22P. An SDS-PAGE band-shift assay shows that PONA is cleaved by *E. coli* Δ2-76 SPase I, but the A22P mutation within PONA prevents cleavage by *E. coli* Δ2-76 SPase I (Fig. S4).

To investigate the binding of *E. coli* SPase I to pre-protein substrate, we utilize surface plasma resonance (SPR) with PONA-A22P bound to the SPR chip and SPase I Δ2-76 or its position 84 mutations were flowed over the SPR chip. The sensorgrams were consistent with the SPase I Pro84 and the mutant P84A having nearly the same steady-state binding affinity (K_D) for the non-cleavable preprotein PONA-A22P whereas the mutants P84L and P84S have approximately 10-fold lower K_D to PONA-A22P, P84 ≅ P84A > P84L ≅ P84S (Table 2, Fig. S5).

Table 2
Binding parameters for *E. coli* SPase I Pro84 Δ2-76 and mutants (P84A, P84L and P84S) to non-cleavable preprotein PONA-A22P.

enzymes	k_a (M^{-1} s^{-1})	k_d (s^{-1})	K_D (M)
E. coli SPase Δ2-76	(5.76 ± 0.37) × 10^3	(4.32 ± 0.42) × 10^{-3}	(7.50 ± 0.72) × 10^{-7}
E. coli SPase Δ2-76 *P84A*	(1.59 ± 0.47) × 10^3	(1.40 ± 0.26) × 10^{-3}	(9.03 ± 0.93) × 10^{-7}
E. coli SPase Δ2-76 *P84L*	(1.03 ± 0.29) × 10^3	(1.40 ± 0.26) × 10^{-3}	(1.41 ± 0.21) × 10^{-6}
E. coli SPase Δ2-76 *P84S*	(5.03 ± 0.78) × 10^2	(8.77 ± 1.37) × 10^{-4}	(1.74 ± 0.03) × 10^{-6}

Parameters represent mean ± SD from three or more experiments.

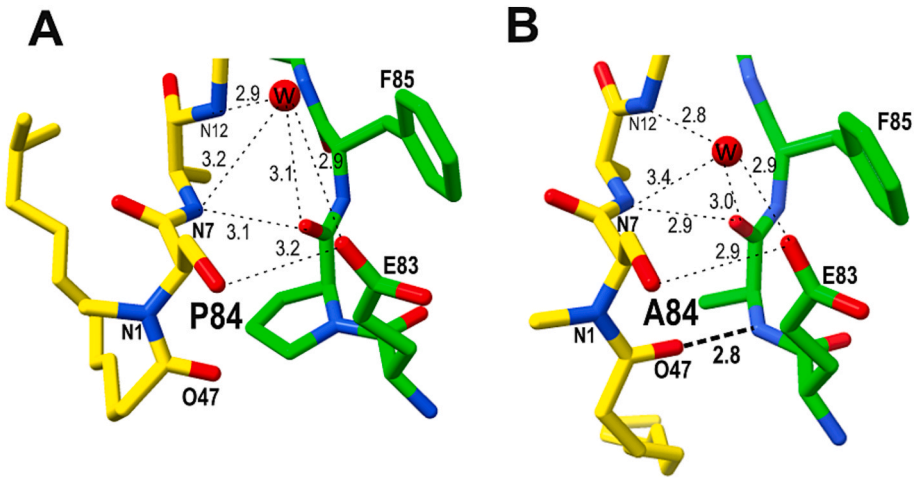


Fig. 4. A close-up view of arylomycin A_2 interactions near residue 84 of *E. coli* SPase I. **A.** *E. coli* SPase I Pro84- arylomycin A_2 complex (PDB ID, 3IIQ). **B.** *E. coli* SPase I P84A – arylomycin A_2 complex (PDB ID, 9NLO). SPase I is shown in stick and colored by element (carbon, green; oxygen, red; nitrogen, blue). The arylomacin A_2 is shown in stick and colored by element (carbon, yellow; nitrogen, blue; oxygen, red). Water molecules are shown as a ball and colored in red. Chain A within the asymmetric unit was used to make the figure. Dashed lines indicate hydrogen bonds and distances are in Ångströms (average value for the two molecules in the asymmetric unit). In figure B, an additional hydrogen bond (2.8Å) that is formed solely due to the introduced mutant P84A is highlighted with thick dashed line.

3.3. Why is proline 84 not conserved in all bacteria?

A comparison of the Pro 84 structure and the P84A structure (Fig. 4) shows that Pro 84 eliminates one hydrogen-bond contact with the arylomycin peptide (proline having no mainchain NH group). The arylomycin binding assay by Smith et al. show that a proline at the 84 position reduced the arylomycin affinity (Smith et al., 2010). Their minimum inhibitory concentration (MIC) assays confirm that having the antibiotic bind less-well is an advantage to the bacteria. It is clear from the work of Smith et al. that a proline at the 84 position results in arylomycin resistance, but it is unlikely that all bacteria have been challenged with arylomycins and this may explain why all bacteria do not have a conserved proline at position 84. It is also unlikely that a proline at the 84-position is always structurally and functionally possible. Future crystal structures of signal peptidase I from a variety of species will be helpful in addressing this issue.

4. Conclusions

The crystal structure of *E. coli* SPase I Δ2-76 P84A in complex with the lipopeptide antibiotic arylomycin A_2 shows that providing a mainchain NH group from a non-proline residue at position 84 contributes to an additional hydrogen bond with a distance of 2.8 Å, consistent with the theory that enzymes with this interaction are susceptible to inhibition by arylomycins.

We asked the question; how is preprotein binding affected by mutations at position 84 of *E. coli* SPase I? The novel preprotein binding assay presented here shows that replacing Pro84 with serine or leucine lowers preprotein substrate binding affinity. This observation is consistent *E. coli* SPase I residue 84 having a role in substrate recognition and binding.

The structure and preprotein binding assay presented here will be helpful in future arylomycin antibiotic development.

CRediT authorship contribution statement

Chuanyun Luo: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation. **Mark Paetzel:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Diffraction data for the research described in this paper was performed using beamline 081D-1 at the Canadian Light Source (CLS), a national research facility of the University of Saskatchewan, which is supported by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council (NSERC), the National Research Council (NRC), the Canadian Institutes of Health Research (CIHR), the Government of Saskatchewan, and the University of Saskatchewan. All other aspects of the research described in this paper was supported through a NSERC Discovery Grant (RGPIN-2023-04229).

Appendix A. . Supplementary material

Supplementary data to this article can be found online at:

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsb.2025.108260.

Data availability

Data will be made available on request.

The structural data for this publication is deposited in the RCSB PDB. PDB reference: *Escherichia coli* signal peptidase I delta 2–76 P84A in complex with lipopeptide inhibitor. Accession number: 9NLO.

References

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., Terwilliger, T.C., 2002. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954.

Adams, P.D., Gopal, K., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Pai, R.K., Read, R.J., Romo, T.D., Sacchettini, J.C., Sauter, N.K., Storoni, L.C., Terwilliger, T.C., 2004. Recent developments in the PHENIX software for automated crystallographic structure determination. J. Synchrotron Radiat. 11, 53–55.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221.

Barkocy-Gallagher, G.A., Bassford Jr., P.J., 1992. Synthesis of precursor maltose-binding protein with proline in the +1 position of the cleavage site interferes with the activity of Escherichia coli signal peptidase I in vivo. J. Biol. Chem. 267, 1231–1238.

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., Leslie, A.G., 2011. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta Crystallogr. D Biol. Crystallogr. 67, 271–281.

Carlos, J.L., Paetzel, M., Brubaker, G., Karla, A., Ashwell, C.M., Lively, M.O., Cao, G., Bullinger, P., Dalbey, R.E., 2000. The role of the membrane-spanning domain of type I signal peptidases in substrate cleavage site selection. J. Biol. Chem. 275, 38813–38822.

Chatterjee, S., Suciu, D., Dalbey, R.E., Kahn, P.C., Inouye, M., 1995. Determination of Km and kcat for signal peptidase I using a full length secretory precursor, pro-OmpA-nuclease A. J Mol Biol 245, 311–314.

Date, T., 1983. Demonstration by a novel genetic technique that leader peptidase is an essential enzyme of Escherichia coli. J. Bacteriol. 154, 76–83.

DeLano, W.L. 2002. The PyMOL Molecular Graphics System.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Evans, P., 2006. Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. 62, 72–82.

Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., 2005. Protein Identification and Analysis Tools on the Expasy Server, p. 571-607, in: J. M. Walker, (Ed.), The proteomics protocols handbook, Humana Press, Totowa, N.J.

Goddard, T.D., Huang, C.C., Meng, E.C., Pettersen, E.F., Couch, G.S., Morris, J.H., Ferrin, T.E., 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Sci. 27, 14–25.

Gouet, P., Robert, X., Courcelle, E., 2003. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 31, 3320–3323.

Holtzel, A., Schmid, D.G., Nicholson, G.J., Stevanovic, S., Schimana, J., Gebhardt, K., Fiedler, H.P., Jung, G., 2002. Arylomycins a and B, new biaryl-bridged lipopeptide antibiotics produced by Streptomyces sp. Tu 6075. II. Structure Elucidation. J Antibiot (tokyo) 55, 571–577.

Kantardjieff, K.A., Rupp, B., 2003. Matthews coefficient probabilities: improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. Protein Sci. 12, 1865–1871.

Kaushik, S., He, H., Dalbey, R.E., 2022. Bacterial Signal Peptides- Navigating the Journey of Proteins. Front. Physiol. 13, 933153.

Kovalevskiy, O., Nicholls, R.A., Long, F., Carlon, A., Murshudov, G.N., 2018. Overview of refinement procedures within REFMAC5: utilizing data from different sources. Acta Crystallogr D Struct Biol 74, 215–227.

Kulanthaivel, P., Kreuzman, A.J., Strege, M.A., Belvo, M.D., Smitka, T.A., Clemens, M., Swartling, J.R., Minton, K.L., Zheng, F., Angleton, E.L., Mullen, D., Jungheim, L.N., Klimkowski, V.J., Nicas, T.I., Thompson, R.C., Peng, S.B., 2004. Novel lipoglycopeptides as inhibitors of bacterial signal peptidase I. J. Biol. Chem. 279, 36250–36258.

Kuo, D.W., Chan, H.K., Wilson, C.J., Griffin, P.R., Williams, H., Knight, W.B., 1993. Escherichia coli leader peptidase: production of an active form lacking a requirement for detergent and development of peptide substrates. Arch. Biochem. Biophys. 303, 274–280.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D.,

Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.
Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallog 26, 283–291.
Liu, J., Luo, C., Smith, P.A., Chin, J.K., Page, M.G., Paetzel, M., Romesberg, F.E., 2011. Synthesis and characterization of the arylomycin lipoglycopeptide antibiotics and the crystallographic analysis of their complex with signal peptidase. J. Am. Chem. Soc. 133, 17869–17877.
Luo, C., Roussel, P., Dreier, J., Page, M.G., Paetzel, M., 2009. Crystallographic analysis of bacterial signal peptidase in ternary complex with arylomycin A2 and a beta-sultam inhibitor. Biochemistry 48, 8976–8984.
Matthews, B.W., 1968. Solvent content of protein crystals. J. Mol. Biol. 33, 491–497.
McCoy, A.J., 2007. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. 63, 32–41.
Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., Ferrin, T.E., 2023. UCSF ChimeraX: Tools for structure building and analysis. Protein Sci. 32, e4792.
Moriarty, N.W., Grosse-Kunstleve, R.W., Adams, P.D., 2009. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Crystallogr. D Biol. Crystallogr. 65, 1074–1080.
Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., Dodson, E.J., 1999. Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr. D Biol. Crystallogr. 55, 247–255.
Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. 67, 355–367.
Nilsson, I., von Heijne, G., 1992. A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a sec-independent protein. FEBS Lett. 299, 243–246.
Paetzel, M., 2014. Structure and mechanism of Escherichia coli type I signal peptidase. BBA 1843, 1497–1508.
Paetzel, M., Goodall, J.J., Kania, M., Dalbey, R.E., Page, M.G., 2004. Crystallographic and biophysical analysis of a bacterial signal peptidase in complex with a lipopeptide-based inhibitor. J. Biol. Chem. 279, 30781–30790.
Paetzel, M., Strynadka, N.C., Tschantz, W.R., Casareno, R., Bullinger, P.R., Dalbey, R.E., 1997. Use of site-directed chemical modification to study an essential lysine in Escherichia coli leader peptidase. J. Biol. Chem. 272, 9994–10003.
Paetzel, M., Chernaia, M., Strynadka, N., Tschantz, W., Cao, G., Dalbey, R.E., James, M. N., 1995. Crystallization of a soluble, catalytically active form of Escherichia coli leader peptidase. Proteins 23, 122–125.
Paetzel, M., Dalbey, R.E., Strynadka, N.C., 1998. Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. Nature 396, 186–190.
Peters, D.S., Romesberg, F.E., Baran, P.S., 2018. Scalable Access to Arylomycins via C-H Functionalization Logic. J. Am. Chem. Soc. 140, 2072–2075.
Pluckthun, A., Knowles, J.R., 1987. The consequences of stepwise deletions from the signal-processing site of beta-lactamase. J. Biol. Chem. 262, 3951–3957.
Roberts, T.C., Smith, P.A., Cirz, R.T., Romesberg, F.E., 2007. Structural and initial biological analysis of synthetic arylomycin A2. J. Am. Chem. Soc. 129, 15830–15838.
Schimana, J., Gebhardt, K., Holtzel, A., Schmid, D.G., Sussmuth, R., Muller, J., Pukall, R., Fiedler, H.P., 2002. Arylomycins a and B, new biaryl-bridged lipopeptide antibiotics produced by Streptomyces sp. Tu 6075. I. Taxonomy, fermentation, isolation and biological activities. J. Antibiot. (Tokyo) 55, 565–570.
Smith, P.A., Roberts, T.C., Romesberg, F.E., 2010. Broad-spectrum antibiotic activity of the arylomycin natural products is masked by natural target mutations. Chem. Biol. 17, 1223–1231.
Smith, P.A., Koehler, M.F.T., Girgis, H.S., Yan, D., Chen, Y., Chen, Y., Crawford, J.J., Durk, M.R., Higuchi, R.I., Kang, J., Murray, J., Paraselli, P., Park, S., Phung, W., Quinn, J.G., Roberts, T.C., Rouge, L., Schwarz, J.B., Skippington, E., Wai, J., Xu, M., Yu, Z., Zhang, H., Tan, M.W., Heise, C.E., 2018. Optimized arylomycins are a new class of Gram-negative antibiotics. Nature 561, 189–194.
Sung, M., Dalbey, R.E., 1992. Identification of potential active-site residues in the Escherichia coli leader peptidase. J. Biol. Chem. 267, 13154–13159.
Tschantz, W.R., Sung, M., Delgado-Partin, V.M., Dalbey, R.E., 1993. A serine and a lysine residue implicated in the catalytic mechanism of the Escherichia coli leader peptidase. J. Biol. Chem. 268, 27349–27354.
Takahara, M., Hibler, D.W., Barr, P.J., Gerlt, J.A., Inouye, M., 1985. The ompA signal peptide directed secretion of Staphylococcal nuclease A by Escherichia coli. J Biol Chem 260, 2670–2674.
Tschantz, W.R., Paetzel, M., Cao, G., Suciu, D., Inouye, M., Dalbey, R.E., 1995. Characterization of a soluble, catalytically active form of Escherichia coli leader peptidase: requirement of detergent or phospholipid for optimal activity. Biochemistry 34, 3935–3941.
Vagin, A.A., Steiner, R.A., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F., Murshudov, G.N., 2004. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr. D Biol. Crystallogr. 60, 2184–2195.
Winn, M.D., Murshudov, G.N., Papiz, M.Z., 2003. Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol. 374, 300–321.