THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 17, pp. 13073–13086, April 27, 2007

Crystal Structure of StaL, a Glycopeptide Antibiotic Sulfotransferase from *Streptomyces toyocaensis*[*] [S]

Received for publication, December 29, 2006, and in revised form, February 20, 2007 Published, JBC Papers in Press, February 28, 2007, DOI 10.1074/jbc.M611912200

Rong Shi[‡1], Sherry S. Lamb[§1,2], Sathesh Bhat[‡], Traian Sulea[¶], Gerard D. Wright[§], Allan Matte[¶], and Miroslaw Cygler[‡¶3]

From the [‡]Department of Biochemistry, McGill University, Montréal, Québec H3G 1Y6, [§]Antimicrobial Research Centre, Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario L8N 3Z5, and [¶]Biotechnology Research Institute, National Research Council of Canada, Montréal, Québec H4P 2R2, Canada

Over the past decade, antimicrobial resistance has emerged as a major public health crisis. Glycopeptide antibiotics such as vancomycin and teicoplanin are clinically important for the treatment of Gram-positive bacterial infections. StaL is a 3′-phosphoadenosine 5′-phosphosulfate-dependent sulfotransferase capable of sulfating the cross-linked heptapeptide substrate both *in vivo* and *in vitro*, yielding the product A47934, a unique teicoplanin-class glycopeptide antibiotic. The sulfonation reaction catalyzed by StaL constitutes the final step in A47934 biosynthesis. Here we report the crystal structure of StaL and its complex with the cofactor product 3′-phosphoadenosine 5′-phosphate. This is only the second prokaryotic sulfotransferase to be structurally characterized. StaL belongs to the large sulfotransferase family and shows higher similarity to cytosolic sulfotransferases (ST) than to the bacterial ST (Stf0). StaL has a novel dimerization motif, different from any other STs that have been structurally characterized. We have also applied molecular modeling to investigate the binding mode of the unique substrate, desulfo-A47934. Based on the structural analysis and modeling results, a series of residues was mutated and kinetically characterized. In addition to the conserved residues (Lys[12], His[67], and Ser[98]), molecular modeling, fluorescence quenching experiments, and mutagenesis studies identified several other residues essential for substrate binding and/or activity, including Trp[34], His[43], Phe[77], Trp[132], and Glu[205].

Over the past decade, antimicrobial resistance has emerged as a major public health crisis. Glycopeptide antibiotics (GPAs)[4] such as vancomycin and teicoplanin are clinically important for the treatment of Gram-positive bacterial infections. Moreover, vancomycin has been commonly termed a "drug of last resort" for treating life-threatening infections by Gram-positive pathogens, many of which are resistant to most other antibiotics (*e.g.* methicillin-resistant *Staphylococcus aureus*) (1, 2). However, the increasing prevalence of infections because of vancomycin-resistant *S. aureus* and *Enterococci* have resulted in elevated rates of morbidity and mortality (3, 4). This serious and growing threat has driven the urgent need for novel antibiotics, including diverse semi-synthetic derivatives of members of this class.

Glycopeptide antibiotics are composed of a heptapeptide core containing both common and unusual amino acids. Based on the identity of the core peptide, GPAs can be divided into two major structural classes exemplified by vancomycin and teicoplanin, respectively. The structural diversity of natural GPAs is mainly derived from enzymatic modifications of the core peptide by a variety of halogenase and glycosyltransferase enzymes (2, 5). *Streptomyces toyocaensis* NRRL 15009 produces A47934 (6, 7), a unique teicoplanin-class glycopeptide because of the following two reasons: (*a*) it is not glycosylated, albeit it contains the characteristic heptapeptide core and is therefore an "aglyco"-glycopeptide antibiotic; and (*b*) it is sulfonated on the N-terminal 4-hydroxyl-L-phenylglycine residue, resulting in its overall negative charge at physiological pH. Naturally occurring aglyco-glycopeptide antibiotics are rare because they are converted to their glycosylated form by glycosyltransferases encoded by the producing organism. However, *S. toyocaensis* lacks glycosyltransferase-encoding genes in the A47934 biosynthetic cluster (8) and as a result exclusively produces A47934, an attractive and proven scaffold for novel antibiotic development both *in vivo* and *in vitro* (9).

Sulfonation of biomolecules has long been known to take place in a variety of organisms, and new biological functions continue to be uncovered in connection with this important transformation (10). Sulfotransferases (ST) are a superfamily of enzymes found in species ranging from bacteria to humans,

* This work was supported in part by the National Research Council of Canada and the Canadian Institutes of Health Research Grants MOP-48370 (to M. C.) and MT-14981 (to G. D. W.) and by a Canada Research Chair in Biochemistry (to G. D. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2OV8, 2OVB, 2OVF) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table 1 and Fig. S1.

[1] Both authors contributed equally to this work.

[2] Supported by the Natural Sciences and Engineering Research Council and by Canadian Institutes of Health Research.

[3] To whom correspondence should be addressed: Biotechnology Research Institute, National Research Council, 6100 Royalmount Ave., Montréal, Québec H4P 2R2, Canada. Tel.: 514-496-6321; Fax: 514-496-5143; E-mail: mirek@bri.nrc.ca.

[4] The abbreviations used are: GPA, glycopeptide antibiotics; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; PAP, 3′-phosphoadenosine 5′-phosphate; ST, sulfotransferase; PDB, Protein Data Bank; BisTris, 2-[bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol; SIE, solvated interaction energy; SeMet, selenomethionine; 5′-PSB, 5′-phosphosulfate-binding motif; r.m.s.d., root-mean-square deviation; mEST, murine estrogen sulfotransferase.

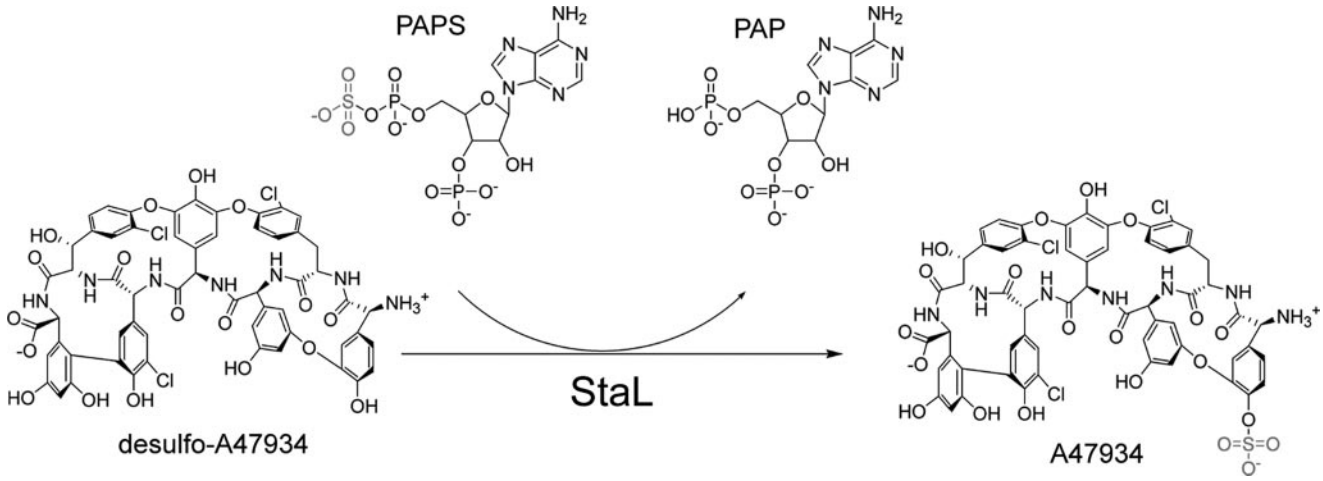


FIGURE 1. **Sulfonation reaction of the glycopeptide antibiotic catalyzed by StaL.** The substrate desulfo-A47934 is sulfonated by StaL using the cofactor PAPS.

catalyzing sulfonation of a variety of endogenous and exogenous substrates. Although the substrate specificity of individual sulfotransferases differs significantly, all utilize the ubiquitous sulfate donor cofactor, 3′-phosphoadenosine 5′-phosphosulfate (PAPS) (10, 11). Studies of this enzyme class have been most extensively performed in eukaryotes (12, 13). The eukaryotic sulfotransferases are grouped into two major classes as follows: cytosolic STs and membrane-associated STs, which share less than 20% sequence identity and differ in terms of their solubility, size, subcellular distribution, and substrate specificity (10–12). The cytosolic STs (30–35 kDa and situated in the cell cytoplasm) catalyze the sulfonation of small molecules, including hormones, bioamines, drugs, and various xenobiotic agents. The membrane-associated STs (45–100 kDa) are located in the Golgi apparatus and have been implicated recently in crucial biological processes. These enzymes sulfonate larger biomolecules, including carbohydrates and proteins (14, 15). Structure-based sequence alignments indicate that both the overall fold and PAPS-binding site are conserved between the two classes. Since the elucidation of the structure of murine estrogen sulfotransferase (mEST) in 1997 (16), the crystal structures of several other, mostly cytosolic, sulfotransferases have been characterized (12).

In contrast to a large body of knowledge accumulated on mammalian sulfotransferases, significantly less is known about the pervasiveness and functions of STs and their sulfated products in bacteria (13). So far, there are only a few bacterial PAPS-dependent sulfotransferases for which the biological significance has been well established. The sulfotransferases NodH and NoeE of *Sinorhizobium meliloti* are essential for establishing symbiosis by sulfonating secreted root lipo-chitooligosaccharide signals, called nodulation factors (17, 18). The enzyme Stf0, a novel mycobacterial sulfotransferase from *Mycobacterium tuberculosis*, has been identified and characterized and is responsible for generating the trehalose 2-sulfate moiety of the putative virulence factor sulfolipid-1 (19). Meanwhile, the crystal structure of Stf0 from *Mycobacterium smegmatis,* the first from a bacterial source, has also been determined. This structure revealed the molecular basis of trehalose recognition and a unique dimer configuration that encloses the substrate within a bipartite active site. Very recently, a sulfotransferase LpsS from *Mesorhizobium loti* has been identified to be involved in sulfonation of cell surface polysaccharides, including lipopolysaccharide and capsule (20). In addition, the sulfotransferase Stf3 has been identified recently, which sulfonates an outer envelope molecule in *M. tuberculosis* involved in mediating virulence in a mouse infection model (21).

StaL (GenBank^{TM} accession number AAM80529) has been annotated as a putative PAPS-dependent sulfotransferase from analysis of the A47934 biosynthetic gene cluster sequence (8). More recent genetic and biochemical studies have confirmed that StaL is a sulfotransferase capable of sulfating the predicted cross-linked heptapeptide substrate both *in vivo* and *in vitro*, yielding the A47934 product. The sulfonation reaction catalyzed by StaL constitutes the final step in A47934 biosynthesis (Fig. 1) (22). In addition, incubation of purified His_6-StaL with various substrates *in vitro* revealed its substrate specificity and yielded two novel sulfo-glycopeptide antibiotics, sulfo-teicoplanin aglycone and sulfo-teicoplanin (22). This sulfonation adds to the expanding repertoire of engineering possibilities for glycopeptide antibiotics with the potential to develop new compounds with sulfated N-terminal 4-hydroxy-L-phenylglycine residues. Here we report the crystal structure of StaL and its complex with the cofactor product PAP. To our knowledge, this is only the second prokaryotic sulfotransferase, following the structure of Stf0, to be structurally characterized. This structure reveals that StaL is most similar to the cytosolic STs from mammalian sources and less similar to bacterial Stf0. StaL adopts a unique dimerization motif, different from that of all other STs for which crystal structures have been solved. We have also applied molecular modeling and site-directed mutagenesis to investigate the binding mode of the unique substrate desulfo-A47934.

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification—The *staL* gene from *S. toyocaensis* was cloned into a modified pET28a vector (EMD Biosciences, Inc.) and expressed in *Escherichia coli* BL21 (DE3) fused to an N-terminal His_6 tag, as described previously (22). The gene was cloned starting from the second ATG and is miss-

ing the first three codons. Therefore, the expressed StaL has 267 amino acids (excluding the His_6 tag) and lacks the N-terminal Met-Asn-Gly sequence. For the production of selenomethionine (SeMet)-labeled protein, the *E. coli* methionine auxotroph strain DL41(DE3) was transformed with the plasmid (23).

An overnight culture of transformed *E. coli* BL21(DE3) or DL41(DE3) was used to inoculate 1 liter of 2YT medium (or LeMaster medium for SeMet-labeled protein) containing 50 μg/ml kanamycin and grown at 37 °C until the absorbance at 600 nm reached 0.6. Protein expression was induced with 1 mM isopropyl 1-thio-β-D-galactopyranoside followed by 16–20 h of incubation at 16 °C. The cells were harvested by centrifugation (4000 × *g*, 4 °C, 25 min) and stored at −20 °C. The cell pellet was resuspended in 40 ml of lysis buffer (50 mM Hepes, pH 7.5, 0.5 M NaCl, 5% (v/v) glycerol, 20 mM imidazole, pH 8, 10 mM β-mercaptoethanol, and 0.5% (v/v) Triton X-100), and the cells were lysed by sonication (eight times for 15 s, with 15 s between bursts). Cell debris were removed by ultracentrifugation (100,000 × *g*, 45 min, 4 °C). The protein supernatant was loaded onto a 2-ml (bed volume) nickel-nitrilotriacetic acid resin (Qiagen) equilibrated with lysis buffer. After application of the protein sample, the column was washed with 40 ml of lysis buffer, followed by 40 ml of washing buffer consisting of 50 mM Hepes buffer, pH 7.5, 0.5 M NaCl, 5% (v/v) glycerol, 20 mM imidazole, pH 8, 10 mM β-mercaptoethanol. StaL was eluted with the above buffer containing 125 mM imidazole, pH 8. The eluted protein was concentrated by ultrafiltration to 12 mg/ml with concomitant exchange of the buffer to 20 mM Hepes, pH 7.5, 0.2 M NaCl, 5% (v/v) glycerol, and 5 mM dithiothreitol. No attempts were made to remove the His_6 tag. SeMet-labeled protein was purified following the same protocol. Gel filtration indicated that the native enzyme formed dimers in solution.

Crystallization—Initial crystallization conditions were determined by the hanging drop vapor diffusion method at 21 °C using screens from Hampton Research (Aliso Viejo, CA) and Qiagen (Mississauga, Canada). The best crystals of native StaL (form 1) were obtained by equilibrating 1 μl of protein (12 mg/ml in buffer containing 20 mM Hepes, pH 7.5, 0.2 M NaCl, 5% (v/v) glycerol, 5 mM dithiothreitol) mixed with 1 μl of reservoir solution (0.1 M NaAc, pH 4.6, and 3.0 M NaCl) over 1.0 ml of reservoir solution. Crystals grew to a size of 0.2 × 0.2 × 0.1 mm^3 in 2 days at 21 °C. Crystals of StaL belong to space group $P6_122$ with unit cell dimensions $a = 86.0$, $c = 164.8$ Å, and contain one molecule in the asymmetric unit with a V_m of 2.73 $Å^3$ Da^{-1}, corresponding to a solvent content of 55% (24). SeMet-labeled protein crystallized under the same conditions. However, two crystal forms appeared in the same drop, with one of these (form 1) isomorphous to the native crystals. The second crystal form (form 2) belongs to space group $P4_12_12$ with unit cell dimensions $a = 89.2$, $c = 117.3$ Å, with one molecule in the asymmetric unit and a V_m of 3.64 $Å^3$ Da^{-1} corresponding to a solvent content of 66%. For data collection crystals from both forms were transferred to Paratone N oil, and the water layer on the surface of the crystal was carefully removed, and the crystal was picked up in a nylon loop and flash-cooled in a nitrogen stream at 100 K (Oxford Cryosystems, Oxford, UK).

We crystallized StaL in the presence of 10 mM PAP (Sigma) from a reservoir solution containing 0.2 M ammonium sulfate, 0.1 M BisTris, pH 5.5, and 22% (v/v) PEG-3350. These crystals (form 3a) belong to the same $P6_122$ space group but are not isomorphous to form 1 crystals and have $a = 88.7$, $c = 169.7$ Å.

An additional crystal form of StaL obtained by co-crystallization with 10 mM PAP was obtained from reservoir solution containing 0.2 M NaI, 0.1 M BisTris propane, pH 6.5, 20% (v/v) PEG-3350. These crystals (form 4) belong to space group $P3_221$ with unit cell parameters $a = 122.3$, $c = 188.2$ Å with four molecules in the asymmetric unit and a V_m of 3.17 $Å^3$ Da^{-1} corresponding to a solvent content of 61%.

Finally, we crystallized StaL (12 mg/ml) in the presence of 10 mM PAP and 5 mM desulfo-A47934 substrate. Hanging drop vapor diffusion over a reservoir solution containing 0.2 M NaI, 0.1 M BisTris propane, pH 6.5, 18% (v/v) PEG-3350 yielded crystals isomorphous to form 3, which we will refer to as form 3b. These crystals grew in 2–4 days and were cryo-protected by a brief transfer to a solution containing reservoir solution supplemented with 15% (v/v) ethylene glycol and flash-cooled in the nitrogen stream at 100 K.

Determination of Product Affinity by Fluorescence Quenching—Steady-state fluorescence spectra were recorded at 25 °C on a Spex Fluorolog 1681 spectrometer (Horiba Jobin Yvon Inc.). A cuvette containing 500 μl of protein sample was prepared at a concentration of 1.5–3 μM in 20 mM Hepes, pH 7.5, 0.2 M NaCl, 5% (v/v) glycerol, and 5 mM dithiothreitol. The intrinsic tryptophan fluorescence of StaL was measured by recording the emission spectra from 300 to 370 nm (2 nm slit width) with a fixed excitation wavelength of 290 nm. Fluorescence quenching experiments were performed by titration of the protein sample with a concentrated stock solution of A47934 dissolved in the same buffer as the protein sample. The concentration of A47934 was varied in the range of 8–135 μM. To eliminate inner filter effects, the absorbance of the enzyme sample did not exceed 0.1 at 290 nM. The spectra were corrected for the intrinsic fluorescence of the A47934 product at each point in the titration.

To analyze the fluorescence data, the fluorescence intensities at various concentrations of the quencher were fitted to the Stern-Volmer Equation 1 for the quenching of emitting fluorophores (25),

$$F_0/F = 1 + K_{SV}[L] \quad \text{(Eq. 1)}$$

where F_o is the fluorescence emission intensity in the absence of quencher; F is the intensity in the presence of quencher at concentration [L]; and K_{SV} is the Stern-Volmer constant (association or affinity constant). A plot of F_o/F *versus* [L] should yield a straight line with a slope of K_{SV} and an intercept of 1.

X-ray Data Collection, Structure Solution, and Refinement—Diffraction data from a form 1 SeMet-labeled StaL crystal were collected to 2.8 Å resolution at three wavelengths (MAD regime) with a Quantum-4 CCD detector (Area Detector Systems Corp., San Diego). Unless otherwise indicated, this and all subsequent data sets were collected at beamline X8C at the National Synchrotron Light Source, Brookhaven National Laboratory. Data integration and scaling was performed with HKL2000 (26). Seven of the nine expected selenium atoms in the asymmetric unit were located with the program SOLVE

(27) using data to 3.0 Å resolution. Solvent flattening resulted in a figure of merit of 0.61. The density-modified phases were utilized with a data set from a native crystal (form 1) collected to 2.58 Å resolution at beamline X25 and extended to the resolution limit. Automatic model building with RESOLVE (28) resulted in a starting model containing ~50% of the main chain atoms and ~20% of side chains. Remaining parts of the model were built manually using the program O (29), and refinement was carried out using the program REFMAC (30). No σ cutoff was used in refinement. In the final rounds of refinement, we applied the translation-libration-screw model for anisotropic temperature factors (31, 32) converging to a final R_{work} and R_{free} of 0.249 and 0.279, respectively. The N-terminal His_6 tag and residues 104–109 and 217–236 were disordered and could not be traced in the electron density map.

A data set for form 3a crystals of StaL crystallized in the presence of PAP and ammonium sulfate was collected to 2.61 Å resolution. The structure was solved by molecular replacement with the program MOLREP (33), using the apo-StaL structure as the search model, and refined to an R_{work} and R_{free} of 0.228 and 0.269, respectively. Only the His_6 tag and residues 217–235 were disordered. No density corresponding to PAP was found in the resulting maps, but instead, two sulfate ions were modeled in the PAP-binding site. We refer to this model as StaL-SO_4.

Crystal form 3b of StaL crystallized in the presence of PAP and desulfo-A47934 yielded data extending to 2.95 Å resolution. The structure was solved by molecular replacement using the apo-StaL model and refined to an R_{work} and R_{free} of 0.224 and 0.269, respectively. The resulting difference electron density clearly showed the presence of a molecule of PAP bound in each subunit, but no density for the antibiotic was found. The final model contained only one disordered segment, consisting of residues 217–235. We refer to this model as StaL-PAP.

The tetragonal form 2 crystals diffracted to 3.45 Å, and the structure was solved by molecular replacement. We identified a dimeric association of the molecules similar to that observed in the hexagonal crystal forms, but the structure was not refined further because of the inferior resolution.

Finally, the trigonal form 4 crystals diffracted to 3.05 Å at beamline X12C. There are four independent molecules, and the model was only partially refined. Electron density indicated the presence of PAP. but it is less well defined than in form 3b crystals. The form 2 and form 4 crystals diffracted to low resolution, showed dimers indistinguishable from those in better diffracting crystal forms, and were not pursued further. Final data collection and refinement statistics are shown in Table 1.

Site-directed Mutagenesis—Oligonucleotide primers were designed using PrimerSelect (DNAstar Inc.) and are listed in supplemental Table 1. Single site mutants were created using the QuikChange® site-directed mutagenesis kit (Stratagene, La Jolla, CA) with modifications as described below. Synthesis of oligonucleotide primers and sequencing of DNA were performed at the MOBIX Lab Central Facility (McMaster University). Using the previously constructed plasmid *staL*/pET28a (22) as a template, PCR conditions were as follows: 5 min at 95 °C, 16 cycles (1 min at 95 °C, 1.5 min at 55 °C, and 10 min at 68 °C), 20 min at 68 °C with 1 unit of *Pfu Ultra*® DNA polymerase (Stratagene), 25 or 50 ng of template DNA, 2.5 mM dNTPS, and ~250 ng of each primer. Subsequent digestion with 10 units of DpnI (New England Biolabs) for 1 h at 37 °C facilitated destruction of the template DNA. Mutagenized plasmids were then transformed into *E. coli* TOP10 (Invitrogen) by electroporation. Plasmids were isolated from the transformants and were analyzed by restriction enzyme digests and DNA sequencing. Plasmids containing the confirmed *staL* point mutations were then transformed into the expression strain *E. coli* BL21(DE3).

Expression, Purification, and Characterization of StaL Mutants—Expression and purification procedures previously reported for wild-type StaL (22) were used for each StaL mutant. Similar expression levels were observed for all mutants except L48K, L48E, and H67A, which exhibited poor expression. Following purification and concentration, *in vitro* reactions were set up for each mutant with 0.1 mM desulfo-A47934 as the substrate. Conditions were as reported previously (22), except that the final concentration of PAPS was 5 mM. Four time points were determined in duplicate for each mutant protein to obtain the initial rate (v_0). Full steady-state kinetic analyses were not performed given the significant substrate inhibition observed by this enzyme (22). We therefore determined only v_0 of the mutants under the substrate conditions indicated. Reactions were analyzed by reverse phase high pressure liquid chromatography using previously reported conditions (22). For each reaction time point, the total peak area was determined for the substrate (*i.e.* desulfo-A47934) and the product (A47934). The relative area of product was then expressed as a fraction of the total substrate present (10 nmol) and plotted *versus* time. The v_0 was determined by linear regression. For comparing the activity of the mutants, the initial rate of each StaL mutant was expressed as a percentage of the wild-type StaL initial rate (Table 2).

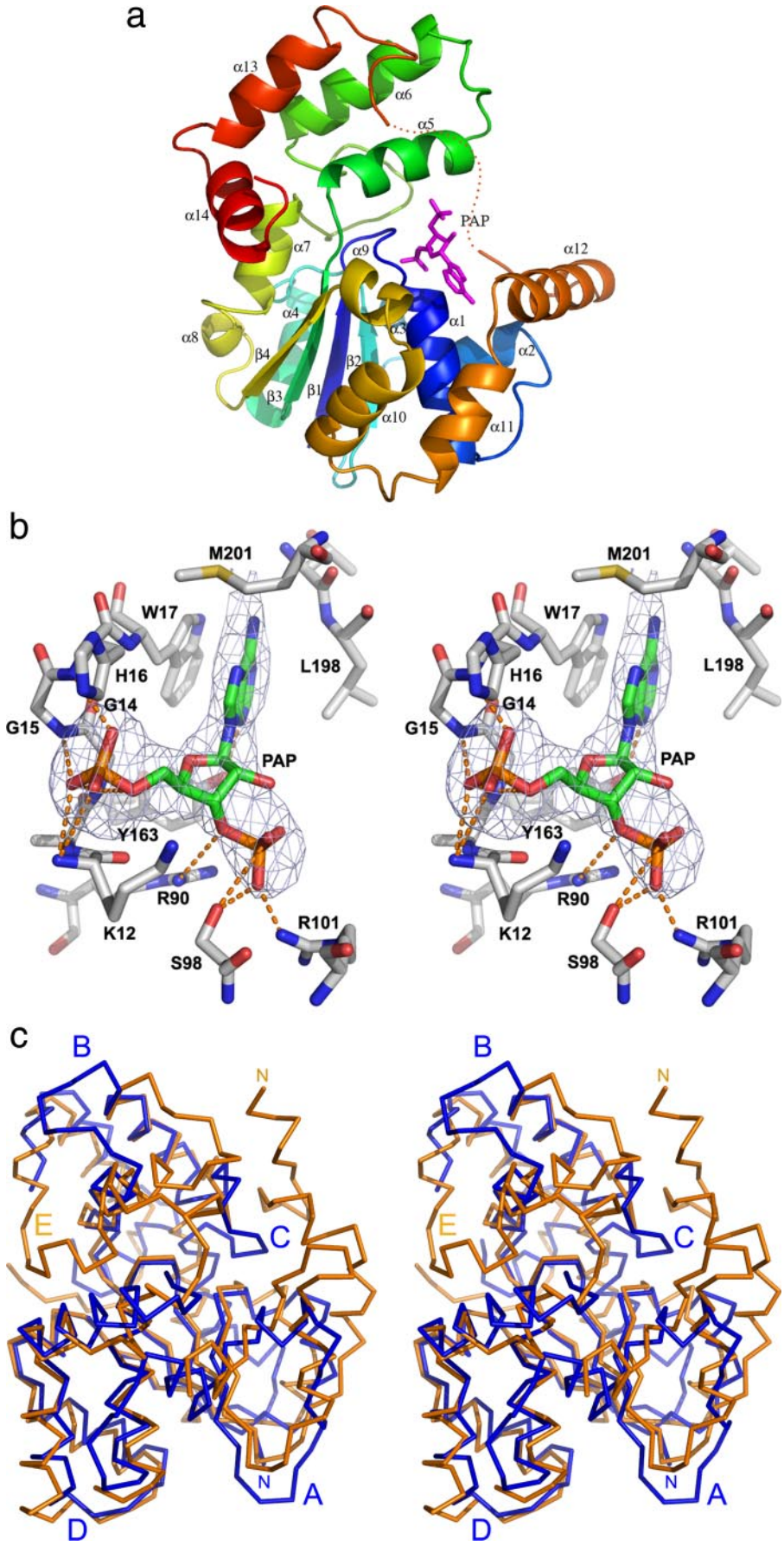
Mutant proteins were analyzed by CD spectroscopy to confirm that mutation did not perturb global protein structure (supplemental Fig. S1). Prior to analysis, the mutant proteins were dialyzed for 3 h at 4 °C in buffer (5 mM Hepes, pH 7.5, and 100 mM NaCl), followed by measurement of the protein concentration by the Bradford assay (34). An AVIV model 215 circular dichroism spectrometer (AVIV Associates, Lakewood, NJ) was used to record the CD spectra. A 1-mm path length quartz cell containing 300 μl of each protein sample (~0.25 mg/ml) was maintained at 25 °C in a thermoelectrically controlled cell holder. Data were collected every 1 nm with an averaging time of 5 s and are expressed as the mean residue ellipticity in units of degrees·cm^2/dmol.

Molecular Modeling of the StaL-PAPS-Desulfo-A47934 Complex—The PAPS cofactor molecule was constructed from the bound PAP, to which a sulfate group was linked in the conformation observed in the human estrogen sulfotransferase-PAPS complex (PDB code 1HY3 (35)). Residues Leu^{213} and Leu^{238} at the ends of the disordered loop 217–235 were blocked; crystallographic waters were removed, and hydrogen atoms were added and oriented to favor hydrogen bonding. The StaL-PAPS complex was conjugate gradient energy-minimized using the AMBER force field (36, 37), an 8 Å nonbonded cutoff, a distance dependent dielectric ($4R_{ij}$), and using AM1-BCC partial charges (38) for PAPS. Harmonic potentials were

TABLE 1
Data collection and refinement statistics

Dataset	SeMet			Apo	Sulfate complex	PAP complex
	Inflection	Peak	Remote			
Unit cell						
Space group	$P6_122$					
a (Å)	86.7			86.0	88.7	87.4
c (Å)	164.0			164.8	169.7	169.4
Resolution (Å)	50-2.82 (2.92-2.82)	50-2.82 (2.92-2.82)	50-2.83 (2.93-2.83)	50-2.58 (2.67-2.58)	50-2.61 (2.70-2.61)	50-2.95 (3.06-2.95)
Wavelength(Å)	0.9795	0.9794	0.9643	1.0	1.1	1.1
Observed *hkl*	116,892	116,683	114,943	357,190	132,667	96,691
Unique *hkl*	16,624[a] (unmerged)	16,599[a] (unmerged)	16,360[a] (unmerged)	12,023	12,609	8,661
Redundancy	7.0	7.0	7.0	29.7	10.5	11.2
Completeness (%)	99.5 (100)	99.5 (100)	99.6 (100)	99.9 (100)	99.3 (93.4)	99.9 (100)
R_{sym} [b]	0.033 (0.512)	0.036 (0.498)	0.032 (0.514)	0.055 (0.471)	0.051 (0.325)	0.048 (0.532)
I/σ (I)	21.2 (4.0)	21.5 (4.1)	23.5 (4.0)	20.7 (8.6)	22.4 (4.2)	17.4 (4.9)
Wilson B (Å^2)				84.7	76.8	101.8
Refinement						
Resolution (Å)				50-2.58	50-2.61	50-2.95
R_{work} (#*hkl*)[c]				0.249 (11,397)	0.228 (11,942)	0.224 (8,187)
R_{free} (#*hkl*)				0.279 (572)	0.269 (612)	0.269 (418)
B-factor (Å^2) (# atoms)						
Protein				77.1 (1,768)	58.8 (1,898)	65.8 (1,873)
Solvent				64.8 (13)	55.6 (33)	59.2 (12)
Ligands					53.3 (15)	79.7 (27)
Ramachandran						
Allowed (%)				98.0	99.5	97.0
Generous (%)				1.0	0.0	2.5
Disallowed(%)				1.0	0.5	0.5
R.M.S.D						
Bonds (Å)				0.014	0.010	0.013
Angles (°)				1.48	1.19	1.61
PDB code				2OV8	2OVB	2OVF

[a] Friedel pairs were unmerged.
[b] $R_{sym} = (\Sigma|I_{obs} - I_{avg}|)/\Sigma I_{avg}$.
[c] $R_{work} = (\Sigma|F_{obs} - F_{calc}|)/\Sigma F_{obs}$.



applied during minimization to constrain various regions of the complex to their crystallographic positions.

The structure of the desulfo-A47934 substrate was built starting from the crystal structure of A-40926 aglycone (Cambridge Crystallographic Data Centre code 134958) (39) that has the same backbone. Only minor modifications were required, *i.e.* desulfo-A47934 lacks the first N-terminal methyl group and otherwise differs only in the positions of three chlorine atoms. The fast rigid exhaustive docking (FRED) program (40) was used to generate and rank initial potential binding modes (poses) of desulfo-A47934 to the StaL-PAPS complex. The top-scored 10,000 substrate poses were filtered down to 500 poses of good shape complementarity and having the reactive hydroxyl oxygen within 4 Å of the sulfur atom of PAPS. All these substrate poses had several minor clashes with StaL, suggesting that some conformational change must occur in the putative binding site to accommodate the substrate. The retained complexes were energy-minimized using the SZYBKI program (OpenEye, Inc., Santa Fe, NM), with the substrate allowed to move as a rigid body, although protein side chains within 10 Å of the substrate and PAPS were flexible. The resulting 500 minimized poses were then scored for binding using our solvated interaction energy (SIE) function as shown in Equation 2,

$$\mathrm{SIE} = E_{\mathrm{inter}}^{\mathrm{VDW}} + E_{\mathrm{inter}}^{\mathrm{Coul}} + E_{\mathrm{inter}}^{\mathrm{HB}} + \Delta G_{\mathrm{solv}}^{\mathrm{elst}} + \Delta G_{\mathrm{solv}}^{\mathrm{np}} - T\Delta S_{\mathrm{bind}}^{\mathrm{rot,trans}} \quad \text{(Eq. 2)}$$

which accounts for intermolecular van der Waals, Coulomb, and hydrogen bond interactions ($E_{\mathrm{inter}}^{\mathrm{VDW}}$, $E_{\mathrm{inter}}^{\mathrm{Coul}}$, and $E_{\mathrm{inter}}^{\mathrm{HB}}$), electrostatic and nonpolar desolvation ($\Delta G_{\mathrm{solv}}^{\mathrm{elst}}$ and $\Delta G_{\mathrm{solv}}^{\mathrm{np}}$), and rotational and translational entropy changes ($-T\Delta S_{\mathrm{bind}}^{\mathrm{rot,trans}}$) upon binding. $E_{\mathrm{inter}}^{\mathrm{VDW}}$ is based on AMBER van der Waals energies scaled by a factor of 0.069. $E_{\mathrm{inter}}^{\mathrm{HB}}$ penalizes deviations from the ideal hydrogen bond geometry and incorporates a scaling factor of 0.8. The electrostatic calculations employed AMBER and AM1-BCC partial charges for protein and ligands, respectively, dielectrics of 20 for solute and ∞ for solvent, and a boundary element solution to the Poisson equation (41). $\Delta G_{\mathrm{solv}}^{\mathrm{np}}$ is derived from a molecular surface area coefficient of 11 cal/(mol·Å^2). The molecular surface was generated with a variable radius probe (42).

The StaL-PAPS-desulfo-A47934 complex with the best SIE-scored substrate pose was further optimized by AMBER energy minimization. The substrate, the cofactor, and the protein atoms around them were allowed to relax under a set of harmonic space constraints, although the distance between the reactive oxygen atom of desulfo-A47934 and the sulfur atom of PAPS was constrained within the 2.0–3.5 Å range. The space constraints were then lifted, and the entire complex was minimized up to a root mean square gradient of 0.001 kcal/(mol·Å).

RESULTS AND DISCUSSION

Overall Structure—We have crystallized StaL using either PEG3350 or sodium chloride at high concentration as a precipitating agent. Under both conditions residues 217–235 were disordered and could not be traced in the electron density map. This segment of polypeptide is rich in Gly residues (6 of 19) and probably requires substrate binding to be ordered. In crystals obtained in the presence of high salt (3 M NaCl, 0.1 M NaAc, pH 4.6), one other region was disordered, residues 104–109.

Each StaL monomer has an α/β architecture. The core of the molecule is a three-layered $\alpha/\beta/\alpha$ sandwich with a central four-stranded parallel β-sheet with the order β2-β1-β3-β4, a bundle of α-helices on one side, and an α-helix and a distorted helical turn on the other side (Fig. 2). The top of the β-sheet is covered by another α-helical bundle. The main features are characteristic for the Rossmann fold with the P-loop crucial for nucleotide phosphate binding. The P-loop following strand β1 contains the sequence Lys-*X*-Gly, where *X* is a residue having a small side chain, which is a hallmark of the PAPS-binding site (12). A large cleft formed between the two helical bundles is readily visible in the molecular surface representation of StaL. A molecule of PAP is located at one end of this cleft, and the other side is wide open and forms the substrate-binding site. The disordered region 217–235 connects the segments located on opposite sides of this cleft and may act as a gate that controls access to the substrate-binding site.

StaL Belongs to a Large Family of Sulfotransferases—Sequence comparisons clearly place StaL in the sulfotransferase family (PFAM PF00685 (43)). These 270–300-amino acid-long proteins or domains are characterized by two sequence motifs involved in PAPS binding. These motifs are also present in StaL, namely the 5′-phosphosulfate-binding motif (5′-PSB) corresponding to the sequence ^{11}P**K**A**G**G**H**^{16} and the 3′-phosphate-binding motif (3′-PB) corresponding to ^{89}I**R**NPRDAML**S**L^{99} (boldface amino acids form direct contacts with PAPS) (44). A search for structural homologs of StaL using DALI (45) reveals its similarity to other sulfotransferases. The closest similarity (*Z* score of 15.8) is to the murine estrogen sulfotransferase SULT1E1 (PDB code 1AQU). Despite the relatively low sequence identity between these two proteins (19%), the root-mean-square deviation (r.m.s.d.) for 204 C-α atom pairs is 2.8 Å (r.m.s.d. for the core 142 C-α atom pairs is 1.5 Å). In general, StaL shows the highest structural similarity to cytosolic eukaryotic sulfotransferases and, somewhat surprisingly, is less similar to the previously solved prokaryotic sulfotransferase Stf0 from *M. smegmatis* (PDB code 1TEX) (19) (only core 88 C-α atom pairs overlap well with r.m.s.d. 1.7 Å) with which it has a sequence identity of less than 15%.

The amino acid sequence of StaL lacks the N-terminal ~30–35 residues that are found in the sequences of some other

FIGURE 2. **Complex of StaL with PAP.** *a,* the schematic representation of StaL complexed with PAP. The protein is rainbow colored (from *blue* at the N terminus to *red* at the C terminus). The PAP molecule is shown in *stick* representation and colored in *magenta*. The location of the disordered loop 217–235 is schematically marked by a *dotted line*. *b,* PAP-binding site in the StaL-PAP complex. Omit electron density for PAP is contoured at 3σ level. The hydrogen bonds between PAP (colors: *C, green*; *O, red*; *N, blue*; *P, orange*) and StaL (colors: *C, white*; *O, red*; *N, blue*; *S, yellow*) are shown by *dashed lines*, *c,* superposition of StaL and murine estrogen ST (PDB code: 1AQU). The StaL and mEST structures are shown in *blue* and *orange*, respectively. *A–D* indicate the regions that differ significantly between the two structures: residues 50–62 (*A*), 104–110 (*B*), 125–137 (*C*), and 180–185 (*D*). *E* shows the loop in mEST covering the substrate-binding site, equivalent of which is disordered in StaL The N termini are marked with *N*. Figure prepared using PyMol.

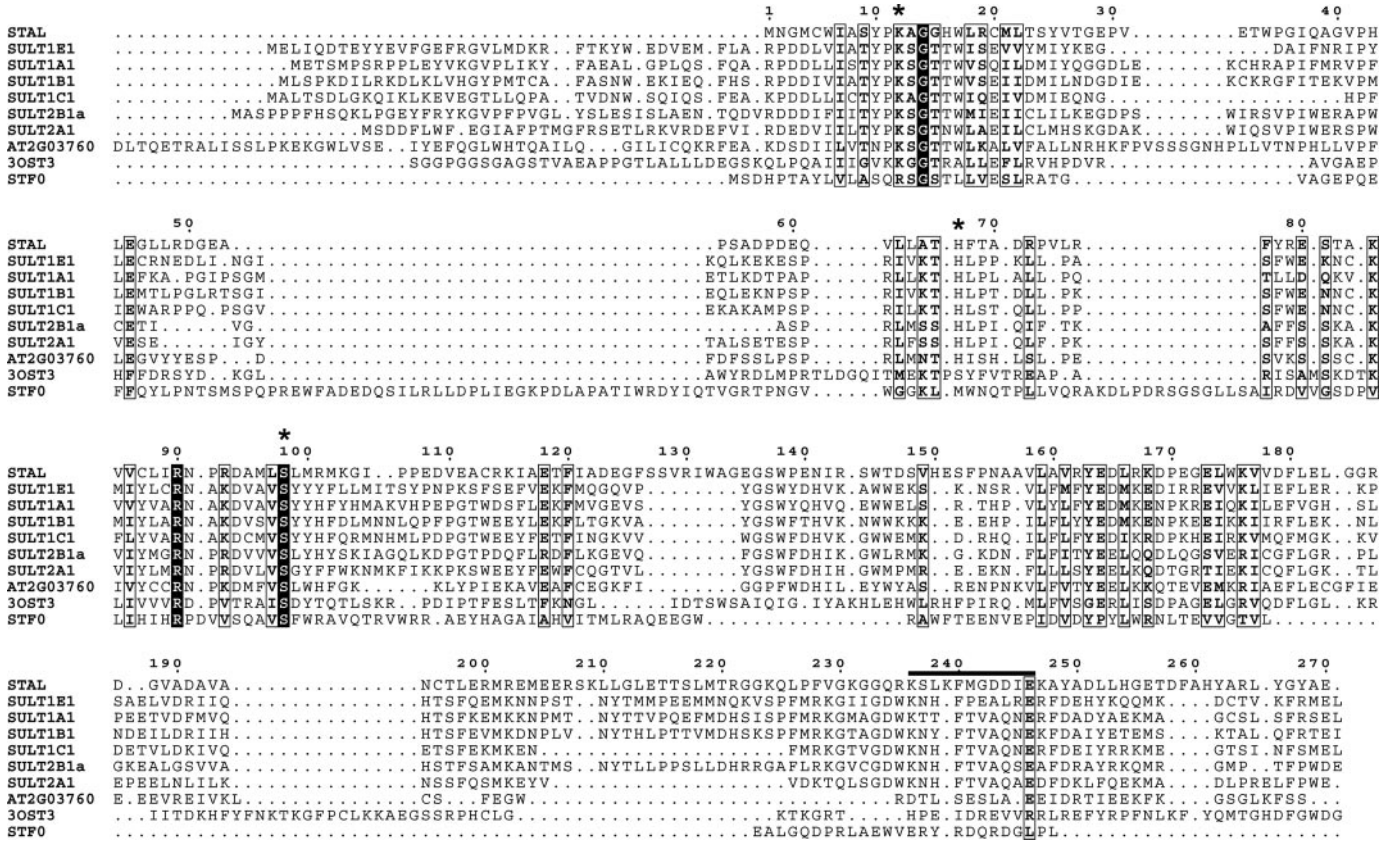


FIGURE 3. **Structure-based sequence alignment.** Sequences of StaL, Stf0 from *M. smegmatis* (PDB code 1TEX (19)), At2g03760 from *Arabidopsis thaliana* (PDB code 1Q44 (54)), SULT1E1 (murine estrogen ST, PDB code 1AQU (16)), and human cytosolic STs SULT1A1 (PDB code 1LS6 (55)), SULT1B1 (PDB code 1XV1 (56)), SULT1C1 (PDB code 2ETG (56)), SULT2A1 (PDB code 1OV4, (57)), SULT2B1a (PDB code 1Q1Q (49)), and human Golgi ST 3OST3 (PDB code 1T8U (50)). *Asterisks* indicate the putative catalytic residues Lys[12], His[67], and Ser[98]. The KTVE motif in cytosolic sulfotransferases is indicated by a *black line* above the sequences. The alignment was obtained with help of sPBDviewer (58) and displayed using program ESPript 2.2.

sulfotransferases (Fig. 2*c* and Fig. 3). In cytosolic STs this region consists of two β-strands and a turn followed by a second α-helix and a turn and covers a hydrophobic patch on the surface of the protein. Some residues in this region are likely part of the substrate binding cavity and contribute to the stable binding of small substrates in cytosolic STs. In StaL, the lack of this region creates a wider substrate-binding site and is in fact a prerequisite for StaL to form its unique dimeric structure, as shown below.

In addition to the N-terminal region, sulfotransferases differ in the length and conformation of several long surface loops. In StaL these regions correspond to residues 50–62 (Fig. 2*c*, *A*), 104–110 (*B*), 125–137 (*C*), and 180–185 (*D*). Part of the segment Leu[44]–Leu[63] protrudes into the region otherwise occupied by the additional N-terminal residues of other cytosolic STs (Fig. 2*c*). The loop Gly[125]–Ser[137] in StaL is longer than the corresponding loop in cytosolic STs and forms one side of the substrate-binding site (Fig. 2*c*, *C*). There is also an insertion of several residues in the region Ser[148]–Ala[157] of StaL relative to other cytosolic STs. However, the most substantial change in StaL is observed in the conformation of the loop encompassing residues Gly[104]–Asp[109], which connects helices α5 and α6. In StaL this loop extends in a direction opposite that seen in other STs relative to the base formed by α5 and α6 helices (Fig. 2*c*, *B*). Although in other STs this loop is directed toward, and contributes to the formation of a compact substrate-binding site, the position of this loop in StaL creates a more open substrate-binding site that is capable of accommodating a large antibiotic as a substrate (Fig. 2*c*). The region 217–235, which we predict covers the substrate-binding site in StaL, is disordered in the absence of substrate, as has also been observed in several other STs where this loop (Fig. 2*c*, *E*) is ordered only when the substrate is bound.

PAPS-binding Site—To identify the PAPS-binding site, we co-crystallized StaL with the PAPS reaction product PAP. We observed clear electron density for a molecule of PAP in crystals obtained from PEG-3350 in the presence of 0.2 M NaI (Fig. 2*b*).

A strand-loop-helix motif containing residues [9]SYPK-AGGH[16] defines the 5′-PSB loop. Hydrogen bonds are formed between the 5′-phosphate and the backbone amide groups of Lys[12], Gly[14], Gly[15], His[16], and the side chain of His[16] (Fig. 2*b*). The conserved residue, Lys[12], located in the vicinity of the 5′-phosphate group, provides charge neutralization of the phosphate group but does not form a direct hydrogen bond with it. Stabilization of the 3′-phosphate of PAP is mainly achieved through a hydrogen bonding interaction with Ser[98]. Mutation of this residue to Ala resulted in a 60% decrease in enzyme activity (Table 2). Two neighboring positively charged residues, Arg[90] and Arg[101], may also help stabilize the 3′-phosphate through hydrogen-bonding interactions. The R101A mutant enzyme showed nearly the same level of activity as the wild-type enzyme (Table 2), suggesting that this residue likely

TABLE 2
Relative activity of StaL mutants

Mutation	Activity[a]
	%
Wild type	100
PAPS-binding mutants	
S98A	40
R101A	92
Antibiotic-binding mutants	
W34F	4
W132F	5
R202A	85
E205A	4
E206A	70
Dimerization mutants	
L48K	ND[b]
L48E	ND
F77E	8
Catalytic mutants	
H43A	18
H67A	ND

[a] Using desulfo-A47934 as a substrate, the initial rate was determined for each mutant and was compared with wild-type StaL.
[b] ND indicates not determined. These mutants did not express well and were not tested.

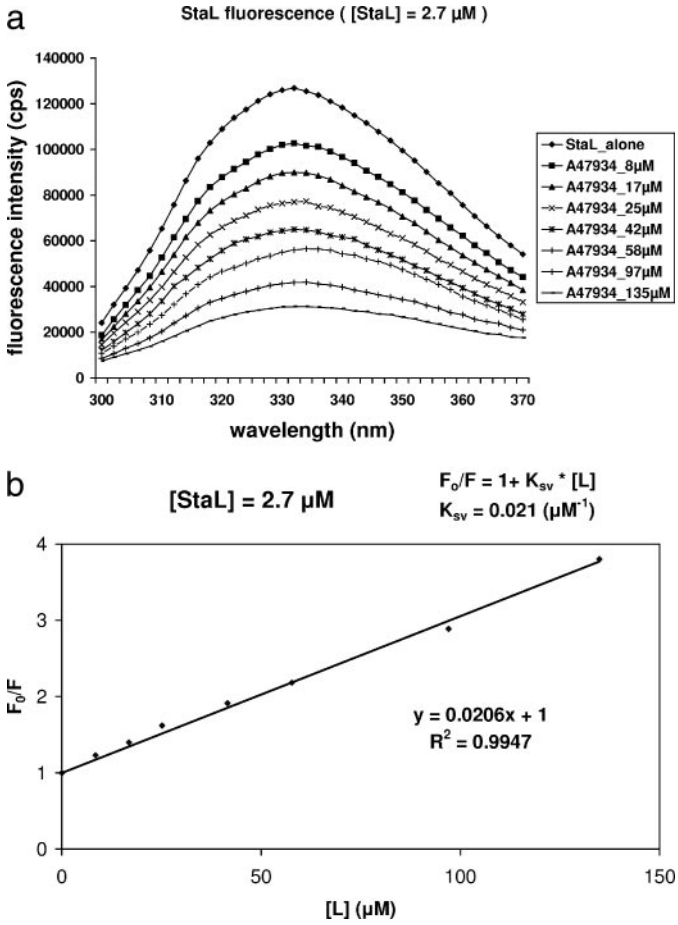


FIGURE 4. **Measurement of product binding by fluorescence.** *a,* steady-state fluorescence emission spectra for StaL. Fluorescence spectra of StaL (concentration 2.7 μM) were obtained in buffer A (20 mM Hepes, pH 7.5, 0.2 M NaCl, 5% (v/v) glycerol, 5 mM dithiothreitol) in the absence or presence of A47934 at 8–135 μM concentration. The spectra were corrected for the intrinsic fluorescence of the A47934 product. *b,* classical Sterm-Volmer plot for A47934 quenching of StaL.

makes a minimal contribution to PAPS binding. The adenine ring contributes to binding by forming a hydrogen bond between its N-3 atom and the hydroxyl group of the highly conserved residue Tyr^{163}. The adenine ring is sandwiched between Trp^{17} and Leu^{198}.

As expected, the PAPS-binding region of StaL is similar to that found in other cytosolic STs, particularly in the 5′-PSB and adenine-binding regions. In all cytosolic STs, the 3′-PB loop (^{257}RKG^{259} in mEST) contributes to the stabilization of the 3′-phosphate. The conserved *GXXGXXK* motif following the 3′-PB loop has been shown to be necessary for substrate binding in murine estrogen sulfotransferase (mEST) (16). This region (Gly^{230}–Lys^{236}) is disordered in StaL in the absence of the substrate, and such disorder has been observed in the structures of several other STs lacking bound ligands. For example, this region is disordered in the crystal structure of the human sulfotransferase SULT1A3 in the absence of substrate (46, 47) but becomes well ordered in the presence of bound substrate (48). In SULT2B1a and SULT2B1b some residues within the N-terminal region, which is part of the substrate-binding pocket and also required for activity, are only ordered upon substrate binding (49). It is therefore likely that the disordered region observed in the structure of StaL complexed with PAP would undergo a disorder-to-order transition upon substrate binding. A unique characteristic of the StaL active site compared with other sulfotransferases is the disulfide bond formed between Cys^{20} from helix α1 and Cys^{196} situated in the turn connecting helices α11 and α12. Other three sulfur-containing residues (Met^{21}, Met^{201}, and Met^{204}) are found to be located in the immediate environment of this disulfide bridge. Both Met^{21} and Met^{201} are within van der Waals contact of Trp^{17} and possibly assist in proper orientation of this side chain for adenine binding. Whether this "cluster" of sulfur atoms is related to the enzymatic properties of StaL remains to be elucidated.

When the crystals were obtained in the presence of 10 mM PAP and 0.2 M $(NH_4)_2SO_4$, we found density only at the expected location of the two phosphates, and we interpreted these peaks as two sulfate ions. This observation suggests that PAP(S)-binding is dependent predominantly on interaction between its phosphate groups and the protein and was outcompeted by the high concentration of sulfate ions. A similar observation was made in the case of sulfotransferase SULT1A3 (46). The binding of PAP causes relatively minor changes in the overall structure of StaL, with an r.m.s.d. of 0.80 Å for all C-α atoms superposing the apo- and PAP-bound forms of StaL. On the other hand, the r.m.s.d. between the StaL-PAP complex and StaL-SO_4 complex (0.48 Å for all C-α atoms) is less, suggesting that the binding of PAP or molecules mimicking the phosphate groups of PAP in the StaL active site induces small but measurable changes in structure.

Fluorescence Quenching of StaL by A47934—StaL was titrated with increasing concentrations of A47934 in the absence of PAP. The tryptophan fluorescence spectrum of StaL was quenched by addition of A47934 (Fig. 4*a*). The change of fluorescence in StaL likely arises from changes in the environment of one or more of the seven tryptophan residues in StaL. Fluorescence titration data fit well the classical Stern-Volmer equation, and the Stern-Volmer constant (K_{SV}) is 0.021 μM^{-1}

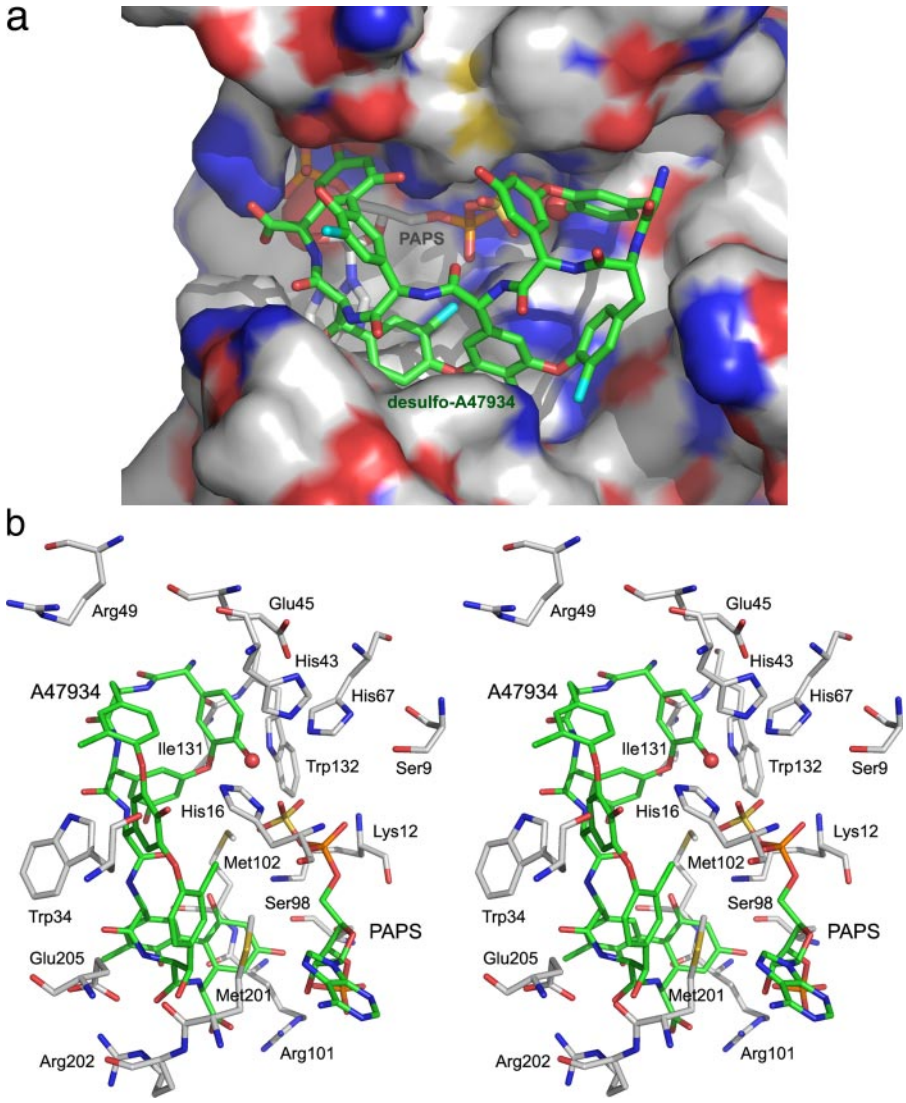


FIGURE 5. **Modeling the binding of desulfo-A47934 and PAPS to StaL.** *a,* surface representation of StaL with the desulfo-A47934 (substrate) and PAPS (cofactor) shown in *stick* representation. The PAPS was modeled on the PAP present in our crystal structure. Desulfo-A47934 was modeled as described in the text (colors: *C, green*; *O, red*; *N, blue*; *Cl, cyan*). The concave surface of desulfo-A47934 packs against the bottom of the binding pocket, and its convex surface is exposed to the solvent but likely covered by the loop that is disordered in our crystal structure. *b,* active site of StaL with modeled desulfo-A47934 and PAPS. The substrate is colored as above, and the protein carbon atoms are painted *white*. The hydroxyl group of desulfo-A47934 to be sulfonated is marked by a *red sphere*.

(Fig. 4*b*), reflecting a moderate affinity (disassociation constant $K_d = 48\ \mu$M) of A47934 for StaL.

A47934-binding Region and Molecular Modeling—Despite several attempts to obtain the structure of the complex between StaL and its substrate desulfo-A47934, either through co-crystallization or by soaking crystals, in the presence or absence of PAP, no density for the antibiotic molecule was observed. We therefore employed molecular modeling to gain insight into the possible mode of substrate binding by StaL.

Protein relaxation by minimization of the molecular mechanics force field-based internal energy was applied both during and after substrate docking, thus ensuring good structural quality of the modeled complex. The protein part of the resulting StaL-PAPS-desulfo-A47934 model is quite similar to the starting crystal structure, with r.m.s.d. values of 0.56, 0.65, and 1.05 Å for the C-α, main chain, and all atoms (including PAP), respectively. These r.m.s.d. figures include relatively larger deviations experienced by only about a half-dozen residues located either at the ends of a disordered loop (217–235) or interacting directly with the substrate. The geometry of the modeled bound substrate, conformationally restricted because of intramolecular cyclization (see Fig. 1), is also close to its initial structure, with an r.m.s.d. of 0.53 Å for all its 83 non-hydrogen atoms. The good quality of the refined model of the complex is reflected quantitatively by its large SIE binding affinity score (see Equation 2) of −12.2 kcal/mol, predicting a dissociation constant of ~1 nM of the desulfo-A47934 substrate from StaL-PAPS. This is a plausible value, given the moderate measured affinity (48 μM) between the A47934 product and apo-StaL derived from fluorescence quenching experiments. A qualitative analysis also indicates that the modeled complex exhibits good intermolecular physicochemical complementarity. The cup-shaped desulfo-A47934 molecule is snugly situated in a pocket, with its concave surface packing against the N terminus of helix α1 (residues 15–20) and the PAPS molecule (Fig. 5*a*). It is stabilized by four adjacent α-helices (α2, α3, α5, and α12) and the loop connecting helices α6 and α7. The convex surface of the substrate molecule is exposed to the solvent, and presumably, the missing loop (residues 217–235) may help to lock the substrate in the pocket.

There are a number of hydrophobic and polar interactions established by the substrate with its putative-binding site (Fig. 5*b*). The stabilization of the substrate in the binding pocket is predicted to be achieved partly through van der Waals contacts with the nonpolar side chains of Trp^{34}, Met^{102}, Ile^{131}, Trp^{132}, and Met^{201}, as well as with aliphatic atoms from the residues Glu^{45}, Arg^{49}, Arg^{101}, and Glu^{205}. Among these, Trp^{34} forms a parallel stacking interaction with the substrate 4-hydroxy-D-phenylglycine moiety at substrate residue 4 (Fig. 5*b*). The W34F mutant enzyme shows a precipitous loss of activity (Table 2), consistent with this model.

Several hydrophilic interactions are also predicted to contribute to substrate binding (Fig. 5*b*). Notably, the reactive phenolic hydroxyl at residue 1 (4-hydroxy-L-phenylglycine) of desulfo-A47934 is well positioned for in-line attack at the sulfate group of PAPS and is hydrogen bonded to the imidazole ring of

His[67], which is predicted to abstract the phenolic proton. Moreover, a hydrogen bonding interaction is established between the amino group at residue 1 of desulfo-A47934 and the main chain carbonyl of Ile[131]. At the other end of the substrate molecule, a hydroxyl group of the 3,5-dihydroxyphenylglycine residue 7 forms hydrogen bonds with both the main chain and side chain of Ser[98] and potentially also with the 3′-phosphate group of PAPS. Mutation of Ser[98] to Ala removes the H-bond donor and results in a 60% drop in activity (Table 2). A salt bridge is established between the carboxylate group of this 3,5-dihydroxyphenylglycine residue and the side chain of Arg[101]. The aliphatic hydroxyl and the main chain amide at substrate residue 6 (D-Tyr) hydrogen bonds to the Glu[205] side chain, whereas the phenolic hydroxyl of residue 4 makes polar contacts with the main chain of Trp[34] and the side chain of His[16]. Mutagenesis studies are consistent with an interaction with this residue as an E205A mutant is >95% impaired (Table 2). Mutation of the adjacent Glu[206] to Ala had a modest impact on activity (30% decrease).

Importantly, docking of the substrate in the modeled binding mode occurs in concert with conformational changes in the side chains of several residues, among which that of Trp[34] is the most notable. Effectively, the Trp[34] side chain swings out of its pocket and is replaced by a segment of the substrate (D-Tyr[6]). Consequently, as mentioned above, a parallel stacking interaction is established between Trp[34] and the substrate. Relatively large structural changes upon substrate binding were also observed at the C terminus of helix α12 preceding the disordered loop 217–235. This indicates that the residues outside this flexible loop are also potentially important for the conformational rearrangement of the substrate-binding site. Importantly, Trp[34] interacts with helix α12 after changing its conformation upon substrate binding. A conformational change involving Trp residues is supported by fluorescence quenching studies of StaL and by site-directed mutagenesis studies that demonstrate that the mutation W34F or W132F results in an enzyme with little activity (Table 2). Although the loop 217–235 of StaL was not modeled, it is expected to be located close to where the substrate docks.

Putative Catalytic Residues—Sulfonation of the phenolic hydroxyl group of the N-terminal 4-hydroxy-L-phenylglycine of the antibiotic requires deprotonation of this group. The pK_a value of the phenolic hydroxyl group is not known; however, it is expected to be similar to Tyr ($pK_a \sim 9-10$). In any event, the enzyme would be required to either bind the deprotonated form of the antibiotic substrate or assist in the removal of the target phenolic hydroxyl. There are two well positioned residues, His[43] and His[67], that are well situated to help stabilize the phenolate negative charge. However, there is likely to be unfavorable charge repulsion by the PAPS sulfate in the substrate-binding pocket if StaL only binds the phenolate form of the substrate. Alternatively, His[43] and His[67] are also well positioned to act in the role of general base, deprotonating the neutral phenol in the active site prior to nucleophilic attack on the PAPS sulfate. The N-ε2 atom of His[67] is closer to the phenolic hydroxyl in the desulfo-A47934-bound model (~3 Å) than in either the apo- or PAP-bound structures (5 Å). Furthermore, Ser[9] is positioned within H-bonding distance of the imidazole ring of His[67] and could facilitate catalysis. Mutation of His[43] impaired but did not abrogate activity (Table 2), whereas the H67A mutant expressed poorly and could not be purified (see below). Given the structural and mutagenesis data, we favor the molecular mechanism of antibiotic sulfation shown in Fig. 6, with His[67] acting as a base in deprotonation of the substrate phenol.

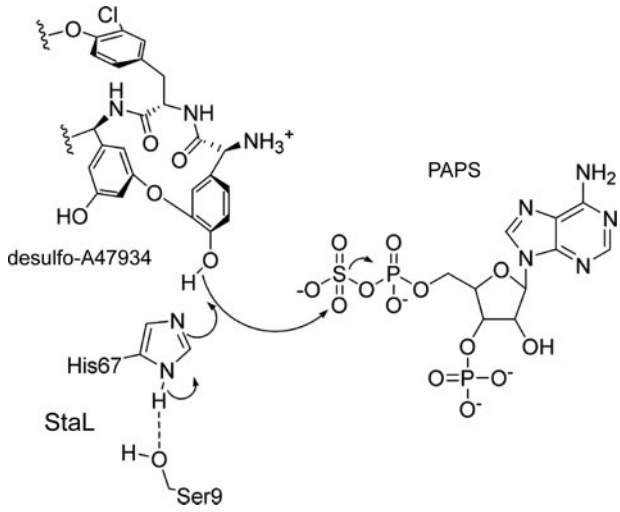


FIGURE 6. **Proposed catalytic mechanism of StaL.** Sulfonation of desulfo-A47934 is assisted by the catalytic base His[67]. The residues Lys[12] and Ser[98] help to position and stabilize the PAPS molecule. Ser[9] is positioned within H-bonding distance of His[67], possibly contributing to catalysis.

Based on the structural similarity of the PAPS-binding site and sequence alignment, the residues Lys[12], His[67], and Ser[98] in StaL (correspond to Lys[48], His[108], and Ser[138] in mEST) are the most likely candidates as catalytic residues (Fig. 5). Structurally equivalent residues have been proposed to be important in catalysis in other sulfotransferases (reviewed in Ref. 11). The lysine and serine residues, which play important roles in positioning and stabilizing the PAPS molecule, are conserved in both cytosolic and Golgi-resident sulfotransferases. It is plausible that these two residues in StaL play similar roles. The structurally equivalent residues in the mycobacterial sulfotransferase Stf0 are Ser[152] and Arg[15]. We propose that the role of a catalytic base that deprotonates the acceptor group is played in StaL by His[67]. This residue is conserved in the cytosolic sulfotransferases and has been postulated to play the same role (11). However, in Golgi resident STs and in *Mycobacterium* Stf0, the corresponding residue is not conserved. Instead, a glutamate at a position equivalent to His[43] in StaL is postulated to play the same role (19, 50). Our attempts to express StaL H67A mutants resulted in insoluble protein preparations, consistent with a role of this residue in the maintenance of a properly folded protein. Thus, the level of structural similarity and the postulated arrangement of catalytic residues in StaL indicate its greater similarity to the cytosolic STs than to the bacterial ST Stf0.

A Novel Mode of Dimerization of STs—Almost all eukaryotic cytosolic sulfotransferases are homodimers in their catalytically active forms, with mEST being one of the few exceptions (10, 51). Structural comparisons and mutagenesis studies have demonstrated that these proteins share a conserved dimeriza-

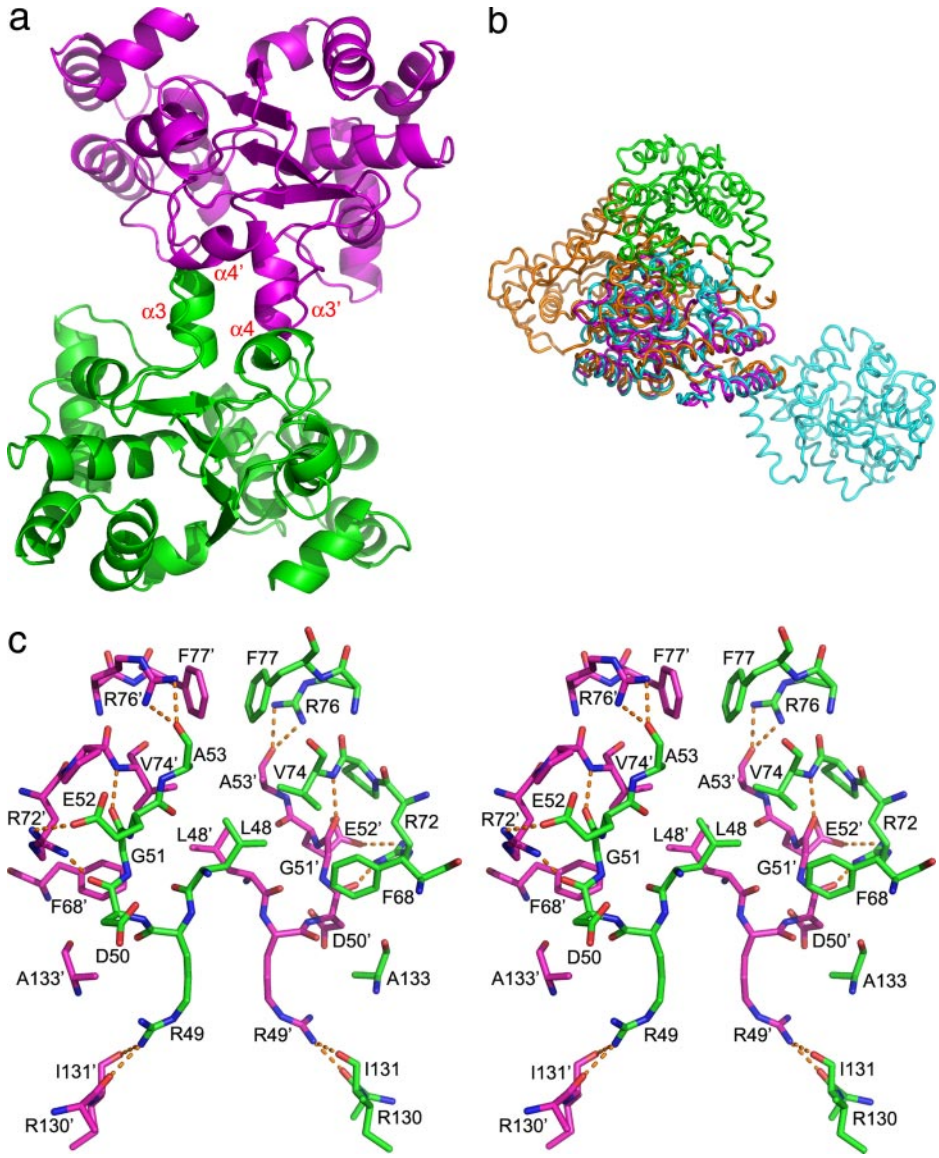


FIGURE 7. **Dimerization interface.** *a,* ribbon diagram of the StaL dimer. The two monomers are colored in *magenta* and *green*, respectively. The dimerization interface mainly involves helices α3 and α4 from each monomer. *b,* superposition of human SULT1A1 dimer (PDB code 1LS6, colored in *cyan*, representative of eukaryotic cytosolic STs), mycobacterial sulfotransferase Stf0 dimer (PDB code 1TEX, colored in *orange*), and StaL dimer colored as in *a*. The dimer interfaces for these STs are located in different regions of the monomers. *c,* residues in the dimerization interface of StaL. The carbon atoms of two monomers are colored in *magenta* and *green*, respectively. The nitrogen and oxygen atoms are shown in *blue* and *red*, respectively. The hydrogen bonds are shown in *dashed lines.*

tion interface represented by the consensus sequence KX_3TVX_3E, designated the KTVE motif, located near the C terminus (51). StaL exists as an apparent dimer as determined by gel filtration chromatography and forms dimers in the crystal structure. Inspection of different pairs of crystallographically related monomers reveals only one possible dimer candidate with a significant interface area, involving residues in helices α3 (^{48}LRDGEA^{53}), α4 (^{72}RPVLRF^{77}), and residues ^{130}RIWA^{133} (Fig. 7). In contrast to a relatively small buried surface area (~3%) for human hydroxysteroid sulfotransferase and other cytosolic STs (51), the buried surface area of the interface in StaL is ~800 Å^2 per monomer, ~6.5% of the total surface. The residues we have identified at the StaL dimerization interface bear no similarity to those of the eukaryotic cytosolic ST dimers (Fig. 7*b*). Nevertheless, it is not unexpected to see a different dimer interface in StaL, considering that StaL does not have the KTVE motif that is conserved in most cytosolic STs. As well, two other crystal forms of StaL belonging to different space groups ($P4_12_12$ or $P3_221$) have the same dimer interface we observe in the $P6_122$ crystal form (data not shown). The prokaryotic sulfotransferase Stf0 also forms dimers but utilizes a surface different from that of StaL and cytosolic STs.

In some sulfotransferase structures, the enzymes crystallize as dimers, although these dimers are unlikely to represent the active form because the putative substrate-binding site is blocked by the dimer interface (46, 52). There is no indication from the structure of the StaL dimer that its dimeric nature would negatively influence access to the substrate, as reflected by our modeling results. Despite the relatively small number of residues (~15) involved in the dimerization interface, the overall interactions are extensive in this region, with more than 10 hydrogen bonds formed between residues of the two monomers (Fig. 7*c*). The formation of the dimer is also driven by numerous hydrophobic contacts. For example, Leu^{48}, situated in the middle part of the interface, interacts with its counterpart residue from the second monomer. Similarly, Phe^{77}, located at the edge of the interface, stacks against the same residue from the second monomer.

Based on analysis and comparison of the StaL structures, two residues located at the dimer interface, Leu^{48} and Phe^{77}, were mutated to assess the contribution of these residues to dimerization. The mutants L48K and L48E resulted in lower protein expression and solubility, which were detrimental to purification of the proteins. These results suggest that Leu^{48} plays an important role in protein stability, which may rely on formation of the dimer. The F77E variant was readily produced and purified using the same protocol as for wild-type StaL, with no significant differences observed in its expression level. Gel filtration experiments showed that it still forms dimers in solution. Nevertheless, its enzymatic activity was significantly reduced, with <10% activity for the substrate desulfo-A47934 compared with wild-type StaL (Table 2). It is evident therefore that Phe^{77}, a residue far

from the active site, is necessary for maximum activity, although it is not essential for dimer formation.

Conclusions—The crystal structure shows that StaL, the final enzyme in the A47934 antibiotic biosynthetic pathway, is more similar to the mammalian cytosolic sulfotransferases than to the bacterial sulfotransferase Stf0. The complex with PAP showed that the cofactor-binding site is similar to that in other STs. However, the substrate-binding site is more open and guarded by a long and flexible loop, which presumably plays a role in holding and release of the substrate. Although we could not obtain the structure of the StaL-PAP-substrate complex, we have used molecular modeling to predict the desulfo-A47934-binding mode and through site-directed mutagenesis confirmed the impact on enzymatic activity of residues predicted to be involved in substrate binding. This enzyme may have potential application in expanding the repertoire of engineering possibilities for developing novel glycopeptide antibiotics.

Acknowledgments—X-ray diffraction data for this study were measured at beamlines X8C, X12C, and X25 of the National Synchrotron Light Source. Financial support for the National Synchrotron Light Source comes principally from the Office of Biological and Environmental Research and of Basic Energy Sciences of the United States Department of Energy, and the National Center for Research Resources of the National Institutes of Health.

REFERENCES

1. Williams, D. H., and Bardsley, B. (1999) *Angew. Chem. Int. Ed. Engl.* **38,** 1172–1193
2. Kahne, D., Leimkuhler, C., Lu, W., and Walsh, C. (2005) *Chem. Rev.* **105,** 425–448
3. Murray, B. E. (2000) *N. Engl. J. Med.* **342,** 710–721
4. Bal, A. M., and Gould, I. M. (2005) *Expert Opin. Pharmacother.* **6,** 2257–2269
5. Nicolaou, K. C., Boddy, C. N., Brase, S., and Winssinger, N. (1999) *Angew. Chem. Int. Ed. Engl.* **38,** 2096–2152
6. Boeck, L. D., and Mertz, F. P. (1986) *J. Antibiot.* (*Tokyo*) **39,** 1533–1540
7. Zmijewski, M. J., Jr., Briggs, B., Logan, R., and Boeck, L. D. (1987) *Antimicrob. Agents Chemother.* **31,** 1497–1501
8. Pootoolal, J., Thomas, M. G., Marshall, C. G., Neu, J. M., Hubbard, B. K., Walsh, C. T., and Wright, G. D. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 8962–8967
9. Solenberg, P. J., Matsushima, P., Stack, D. R., Wilkie, S. C., Thompson, R. C., and Baltz, R. H. (1997) *Chem. Biol.* **4,** 195–202
10. Chapman, E., Best, M. D., Hanson, S. R., and Wong, C. H. (2004) *Angew. Chem. Int. Ed. Engl.* **43,** 3526–3548
11. Negishi, M., Pedersen, L. G., Petrotchenko, E., Shevtsov, S., Gorokhov, A., Kakuta, Y., and Pedersen, L. C. (2001) *Arch. Biochem. Biophys.* **390,** 149–157
12. Rath, V. L., Verdugo, D., and Hemmerich, S. (2004) *Drug Discov. Today* **9,** 1003–1011
13. Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., and Bertozzi, C. R. (2002) *Chem. Biol.* **9,** 767–776
14. Capila, I., and Linhardt, R. J. (2002) *Angew. Chem Int. Ed. Engl.* **41,** 391–412
15. Farzan, M., Mirzabekov, T., Kolchinsky, P., Wyatt, R., Cayabyab, M., Gerard, N. P., Gerard, C., Sodroski, J., and Choe, H. (1999) *Cell* **96,** 667–676
16. Kakuta, Y., Pedersen, L. G., Carter, C. W., Negishi, M., and Pedersen, L. C. (1997) *Nat. Struct. Biol.* **4,** 904–908
17. Roche, P., Debelle, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Denarie, J., and Prome, J. C. (1991) *Cell* **67,** 1131–1143
18. Hanin, M., Jabbouri, S., Quesada-Vincens, D., Freiberg, C., Perret, X., Prome, J. C., Broughton, W. J., and Fellay, R. (1997) *Mol. Microbiol.* **24,** 1119–1129
19. Mougous, J. D., Petzold, C. J., Senaratne, R. H., Lee, D. H., Akey, D. L., Lin, F. L., Munchel, S. E., Pratt, M. R., Riley, L. W., Leary, J. A., Berger, J. M., and Bertozzi, C. R. (2004) *Nat. Struct. Mol. Biol.* **11,** 721–729
20. Townsend, G. E. I., Forsberg, L. S., and Keating, D. H. (2006) *J. Bacteriol.* **188,** 8560–8572
21. Mougous, J. D., Senaratne, R. H., Petzold, C. J., Jain, M., Lee, D. H., Schelle, M. W., Leavell, M. D., Cox, J. S., Leary, J. A., Riley, L. W., and Bertozzi, C. R. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 4258–4263
22. Lamb, S. S., Patel, T., Koteva, K. P., and Wright, G. D. (2006) *Chem. Biol.* **13,** 171–181
23. Hendrickson, W. A., Horton, J. R., and LeMaster, D. M. (1990) *EMBO J.* **9,** 1665–1672
24. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
25. Lehrer, S. S., and Leavis, P. C. (1978) *Methods Enzymol.* **49,** 222–236
26. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
27. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 849–861
28. Terwilliger, T. C. (2002) *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **58,** 1937–1940
29. Jones, T. A., Zhou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
30. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
31. Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003) *Methods Enzymol.* **374,** 300–321
32. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 122–133
33. Vagin, A., and Teplyakov, A. (1997) *J. Appl. Crystallogr.* **30,** 1022–1025
34. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
35. Pedersen, L. C., Petrotchenko, E., Shevtsov, S., and Negishi, M. (2002) *J. Biol. Chem.* **277,** 17928–17932
36. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc.* **117,** 5179–5197
37. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) *J. Comput. Chem.* **25,** 1157–1174
38. Jakalian, A., Jack, D. B., and Bayly, C. I. (2002) *J. Comput. Chem.* **23,** 1623–1641
39. Schäfer, M., Pohl, E., Schmidt-Bäse, K., Sheldrick, G. M., Hermann, R., Malabarba, A., Nebuloni, M., and Pelizzi, G. (1996) *Helv. Chim. Acta* **79,** 1916–1924
40. McGann, M. R., Almond, H. R., Nicholls, A., Grant, J. A., and Brown, F. K. (2003) *Biopolymers* **68,** 76–90
41. Purisima, E. O. (1998) *J. Comput. Chem.* **19,** 1494–1504
42. Bhat, S., and Purisima, E. O. (2006) *Proteins* **62,** 244–261
43. Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2006) *Nucleic Acids Res.* **34,** D247–D251
44. Kakuta, Y., Pedersen, L. G., Pedersen, L. C., and Negishi, M. (1998) *Trends Biochem. Sci.* **23,** 129–130
45. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480
46. Dajani, R., Cleasby, A., Neu, M., Wonacott, A. J., Jhoti, H., Hood, A. M., Modi, S., Hersey, A., Taskinen, J., Cooke, R. M., Manchee, G. R., and Coughtrie, M. W. (1999) *J. Biol. Chem.* **274,** 37862–37868
47. Bidwell, L. M., McManus, M. E., Gaedigk, A., Kakuta, Y., Negishi, M., Pedersen, L., and Martin, J. L. (1999) *J. Mol. Biol.* **293,** 521–530
48. Lu, J. H., Li, H. T., Liu, M. C., Zhang, J. P., Li, M., An, X. M., and Chang, W. R. (2005) *Biochem. Biophys. Res. Commun.* **335,** 417–423
49. Lee, K. A., Fuda, H., Lee, Y. C., Negishi, M., Strott, C. A., and Pedersen, L. C. (2003) *J. Biol. Chem.* **278,** 44593–44599
50. Moon, A. F., Edavettal, S. C., Krahn, J. M., Munoz, E. M., Negishi, M., Linhardt, R. J., Liu, J., and Pedersen, L. C. (2004) *J. Biol. Chem.* **279,** 45185–45193
51. Petrotchenko, E. V., Pedersen, L. C., Borchers, C. H., Tomer, K. B., and Negishi, M. (2001) *FEBS Lett.* **490,** 39–43

52. Pedersen, L. C., Petrotchenko, E. V., and Negishi, M. (2000) *FEBS Lett.* **475,** 61–64
53. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28,** 235–242
54. Smith, D. W., Johnson, K. A., Bingman, C. A., Aceti, D. J., Blommel, P. G., Wrobel, R. L., Frederick, R. O., Zhao, Q., Sreenath, H., Fox, B. G., Volkman, B. F., Jeon, W. B., Newman, C. S., Ulrich, E. L., Hegeman, A. D., Kimball, T., Thao, S., Sussman, M. R., Markley, J. L., and Phillips, G. N., Jr. (2004) *Proteins* **57,** 854–857
55. Gamage, N. U., Duggleby, R. G., Barnett, A. C., Tresillian, M., Latham, C. F., Liyou, N. E., McManus, M. E., and Martin, J. L. (2003) *J. Biol. Chem.* **278,** 7655–7662
56. Dombrovski, L., Dong, A., Bochkarev, A., and Plotnikov, A. N. (2006) *Proteins* **64,** 1091–1094
57. Chang, H. J., Shi, R., Rehse, P., and Lin, S. X. (2004) *J. Biol. Chem.* **279,** 2689–2696
58. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18,** 2714–2723

Crystal Structure of StaL, a Glycopeptide Antibiotic Sulfotransferase from *Streptomyces toyocaensis*

Rong Shi, Sherry S. Lamb, Sathesh Bhat, Traian Sulea, Gerard D. Wright, Allan Matte and Miroslaw Cygler

J. Biol. Chem. 2007, 282:13073-13086.
doi: 10.1074/jbc.M611912200 originally published online February 28, 2007
