THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 9, Issue of February 28, pp. 7655–7662, 2003
© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

Structure of a Human Carcinogen-converting Enzyme, SULT1A1

STRUCTURAL AND KINETIC IMPLICATIONS OF SUBSTRATE INHIBITION*

Received for publication, July 19, 2002, and in revised form, December 4, 2002
Published, JBC Papers in Press, December 5, 2002, DOI 10.1074/jbc.M207246200

Niranjali U. Gamage‡, Ronald G. Duggleby§, Amanda C. Barnett‡, Michael Tresillian‡, Catherine F. Latham‡, Nancy E. Liyou‡, Michael E. McManus‡, and Jennifer L. Martin¶‖

From the ‡Department of Physiology and Pharmacology, School of Biomedical Sciences, §Department of Biochemistry and Molecular Biology, School of Molecular and Microbial Sciences, and ¶Centre for Drug Design and Development and Special Research Centre for Functional and Applied Genomics, Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia

Sulfonation catalyzed by sulfotransferase enzymes plays an important role in chemical defense mechanisms against various xenobiotics but also bioactivates carcinogens. A major human sulfotransferase, SULT1A1, metabolizes and/or bioactivates many endogenous compounds and is implicated in a range of cancers because of its ability to modify diverse promutagen and procarcinogen xenobiotics. The crystal structure of human SULT1A1 reported here is the first sulfotransferase structure complexed with a xenobiotic substrate. An unexpected finding is that the enzyme accommodates not one but two molecules of the xenobiotic model substrate *p*-nitrophenol in the active site. This result is supported by kinetic data for SULT1A1 that show substrate inhibition for this small xenobiotic. The extended active site of SULT1A1 is consistent with binding of diiodothyronine but cannot easily accommodate β-estradiol, although both are known substrates. This observation, together with evidence for a disorder-order transition in SULT1A1, suggests that the active site is flexible and can adapt its architecture to accept diverse hydrophobic substrates with varying sizes, shapes and flexibility. Thus the crystal structure of SULT1A1 provides the molecular basis for substrate inhibition and reveals the first clues as to how the enzyme sulfonates a wide variety of lipophilic compounds.

Sulfonation is a widely distributed biological reaction catalyzed by members of the supergene family of enzymes called sulfotransferases (SULTs).[1] SULTs utilize the sulfonate donor 3′-phosphoadenosine 5′-phosphosulfate (PAPS) to catalyze sulfonation, yielding PAP as the desulfonated product of the reaction. These enzymes modify the biological activities of a diverse range of compounds including neurotransmitters, hormones, and drugs. The reaction aids excretion of foreign chemicals but, in some cases, causes bioactivation of carcinogens and mutagens (1–3). SULTs are the focus of intense research in the fields of cancer, drug metabolism, and pharmacogenetics because genetic variation, particularly in isoenzyme SULT1A1, may predispose to lung cancers (4), protect against colorectal cancers (5), and affect the age of onset in breast cancer (6).

To date, five distinct mammalian gene families for cytosolic SULTs (SULTs 1–5) have been identified (7), although SULT3 and SULT5 have not been found in humans (8). The most extensive group of the human cytosolic SULTs is the SULT1 family, which includes SULTs 1A1, 1A2, 1A3, 1B1, 1C1, 1C2, and 1E1. Each SULT has distinct substrate preferences but may also exhibit broad and overlapping substrate specificity to span the diversity of chemicals requiring sulfonation. For example, SULT1A3 catalyzes dopamine sulfonation but also accepts *p*-nitrophenol (*p*NP) with lower affinity (9, 10). SULT1A1, which shares 93% identity with SULT1A3, also sulfonates dopamine and *p*NP but with the reverse preference, and in addition it can sulfonate a wide range of hydrophobic molecules (Fig. 1) including xenobiotics and endogenous substrates 3,3′-diiodothyronine (T2) (11, 12) and 17β-estradiol (E2) (3, 13).

SULT crystal structures have been determined for mouse estrogen SULT1E1 complexed with PAP/E2 (Protein Data Bank (PDB) codes in parentheses) (1AQU), PAP (1AQY), or PAP/vanadate (1BO6) (14); human estrogen SULT1E1 complexed with PAPS (1HY3) (15); human dopamine SULT1A3 complexed with SO_4^{2-} (1CJM) (16) or PAP (10); human hydroxysteroid SULT2A3 complexed with PAP (1EFH) (17); the SULT domain of heparin sulfate-*N*-deacetylase complexed with PAP (1NST) (18); and human dehydroepiandrosterone sulfotransferase (DEA-SULT) complexed with substrate (1J99) (19). All of these structures incorporate a common PAPS binding region and a variable substrate binding region. Of these, only the mouse SULT1E1-PAP-E2 and the DEA-SULT structures were determined in the presence of an acceptor substrate.

To gain an understanding of the substrate specificity of the human SULT1A family, we have investigated the structures of SULT1A1 and SULT1A3. The similarities and differences between SULT1A family members represent a good experimental model with which to unravel the critical features that determine substrate preferences of cytosolic SULTs. Our approach has involved the use of site-directed mutagenesis and x-ray crystallographic studies of both SULT1A1 and SULT1A3. We reported the first crystal structure of human SULT1A3 (16). Although the substrate was not bound in the structure, with the aid of site-directed mutagenesis and molecular modeling, we and others have identified some of the residues involved in substrate specificity of SULT1A3 (9, 10, 20).

We now report the crystal structure of human SULT1A1

* This project was supported by a grant from the Australian National Health and Medical Research Council. 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 1LS6) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‖ Recipient of an Australian Research Council senior research fellowship. To whom correspondence should be addressed. Tel.: 61-7-3365-4942; Fax: 61-7-3365-1990; E-mail: J.Martin@imb.uq.edu.au.

[1] The abbreviations used are: SULT, sulfotransferase; DEA, dehydroepiandrosterone; E2, 17β-estradiol; PAPS, 3′-phosphoadenosine 5′-phosphosulfate; PAP, 3′-phosphoadenosine 5′-phosphate; PDB, Protein Data Bank; PSB, phosphosulfate binding; *p*NP, *p*-nitrophenol; T2, 3,3′-diiodothyronine; CNS, crystallography NMR sofware.

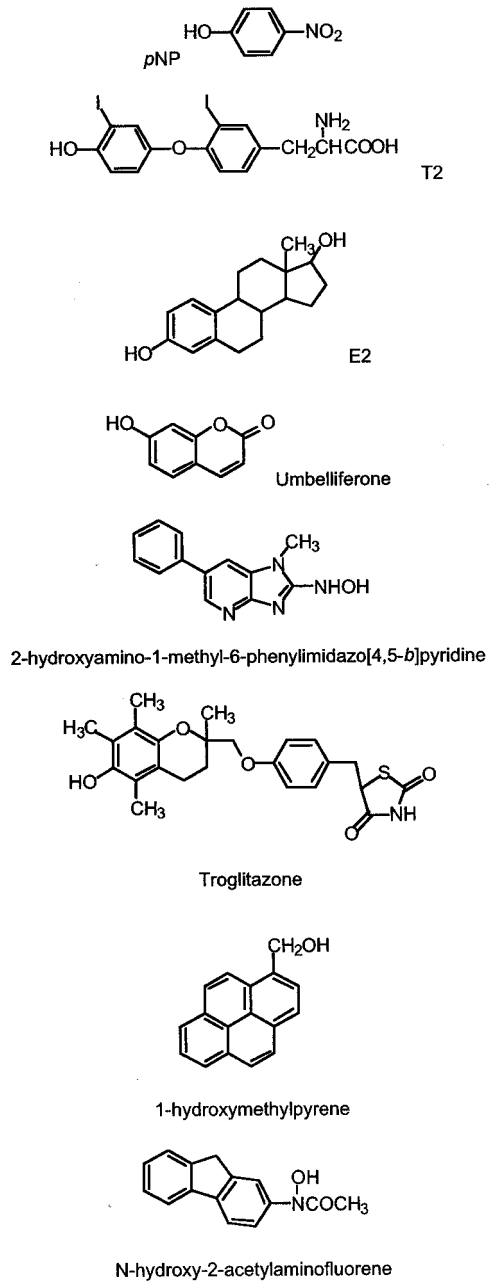


FIG. 1. **Selection of compounds that are sulfonated by SULT1A1 (3, 8, 9, 11, 12, 13, 37).**

crystallized in the presence of PAP and *p*NP. This is the first SULT crystal structure complexed with a xenobiotic substrate, and unexpectedly it reveals two *p*NP molecules in the active site. We show that at low concentrations of this substrate, the kinetics exhibit slight nonhyperbolic behavior and that at high concentrations the enzyme activity is inhibited. These kinetic features are fully consistent with the structure that we have determined. We also show that the endogenous ligand T2 can bind within the substrate binding site in a flexed conformation but that E2 cannot be as easily accommodated.

EXPERIMENTAL PROCEDURES

Expression, Purification, Enzyme Activity, and Crystallization—Human SULT1A1 cDNA incorporating an N-terminal hexahistidine tag was expressed in *Escherichia coli* using a pET-28a(+) vector as described previously (9, 16, 21). SULT1A1 was purified by TALON cobalt affinity (Clontech Laboratories) and gel filtration chromatography and then concentrated to 25 mg ml^{-1} in 20 mM Tris, pH 8.0, and 10 mM dithiothreitol. Protein purity was estimated to be above 95% by SDS-PAGE.

Enzyme activity was determined by the method described by Foldes and Meek (22). The reaction mixture contained 10 mM potassium phosphate, pH 7.0, 20 μM [^{35}S]PAPS, 0.1 μg/ml SULT1A1, and varying concentrations of the substrate (*p*NP 0.1–20 μM) in a final volume of 500 μl. Reactions were initiated by the addition of enzyme to the reaction mixture, which was then incubated for 20 min at 37 °C. The reaction was terminated by precipitation of enzyme and [^{35}S]PAPS with 0.1 M barium acetate, 0.1 M barium hydroxide, and 0.1 M zinc sulfate. The incubation time and protein concentration used were chosen to be within the linear range for product formation using *p*NP as the substrate. Assays were performed in duplicate and corrected for background activity using a control with no substrate added. Radioactivity of the sulfonated *p*NP was quantified using a liquid scintillation counter (Tri-Carb 2500, Packard).

Crystallization was carried out using hanging drop vapor diffusion. Initial conditions were identified from Hampton Research commercial screens and then optimized by incremental scanning around the initial conditions. The final crystallization conditions are as follows. The purified protein was preincubated with 45 mM PAP and 5 mM *p*NP on ice for 30 min. Thereafter, 1 μl of reservoir solution (0.1 M Tris-HCl, pH 8.0, 20% polyethylene glycol 4000) was mixed with 1 μl of protein solution on the coverslip, placed over the well containing the reservoir solution, and incubated at 20 °C. Orthorhombic crystals (size $\sim 0.5 \times 0.3 \times 0.2$ mm) appeared 3–4 days after microseeding in 0.1 M Tris-HCl, pH 8.0, 20% polyethylene glycol 4000.

Diffraction Data Measurement—X-ray diffraction data were measured on an R-AXIS IV++/RU-200 system with Osmic Blue mirror optics. Crystals were mounted in quartz capillary tubes, and diffraction data to 1.9 Å resolution were measured at 17 °C. An oscillation range of 1°, an exposure time of 3 min, and a crystal-to-detector distance of 100 mm were used. The data were integrated, merged, and scaled using the software package Crystal Clear© (Rigaku Corp.).

Structure Determination and Refinement—Monomer A of mouse SULT1E1 (48% sequence identity, PDB code 1AQU (14)) was used as the search model for phasing by molecular replacement. The model was generated by changing to alanine all residues with sequence differences between SULT1A1 and mouse SULT1E1. The model included residues 7–64, 75–242, and 244–295 of mouse SULT 1E1. Rigid body refinement of the molecular replacement solution in CNS software package (23) gave *R*-factor and *R*-free values of 48.9 and 49.2%, respectively. Simulated annealing, by Cartesian molecular dynamics, followed by restrained B-factor refinement improved these values to 35.8% and 38.7%, respectively. Several rounds of model building in O (24) and refinement in CNS software package gave a final *R*-factor of 18.2% and *R*-free of 20.5%. The final model includes all except the first seven residues at the N terminus (there is no density for the additional hexahistidine tag residues). The crystal structure includes bound PAP, two *p*NP molecules, and 154 water molecules. Alternate conformations are modeled for Glu^{13}, Gln^{35}, Val^{54}, Met^{77}, Asp^{156}, Glu^{174}, Glu^{246}, and Met^{260}. The quality of the structure was assessed with PROCHECK (25) and WHATCHECK (26). Coordinates and structure factors have been deposited in the PDB with code 1LS6 (27).

Molecular Modeling—For structural comparison with SULT1A1, the crystal structures of SULT1A3 (PDB code 1CJM (16)) and SULT1E1 (1AQU (14)) were used. For T2 (diiodothyronine) and E2 docking, PAPS was modeled into the SULT1A1 structure based on the SULT1E1-PAPS structure (PDB code 1HY3 (15)) and water and *p*NP molecules were removed from the substrate binding site. T2 (or E2) was docked using GOLD (28) with 10 genetic algorithm runs and default parameters. E2 was also modeled into SULT1A1 using the bound conformation of the ligand in SULT1E1 (PDB code 1AQU). The ligand was oriented in the SULT1A1 active site after superimposing the two protein structures based on a structural alignment of active site residues within 5 Å of the ligand. Solvent-accessible (Connolly) surfaces were calculated in InsightII using a radius of 1.4 Å. Figures were prepared using Molscript (29), Raster3d (30), and InsightII (Accelrys).

RESULTS

The crystal structure of SULT1A1 (Fig. 2*a*) has been refined to 1.9 Å (Table I) with excellent crystallographic and geometric statistics. The structure incorporates the core α/β domain found in all SULTs that forms a central 5-stranded parallel β-sheet surrounded on either side by helices. The SULT1A1 crystal structure is well ordered, with only the first 7 of 295 (2%) residues disordered. This is in stark contrast to the crystal

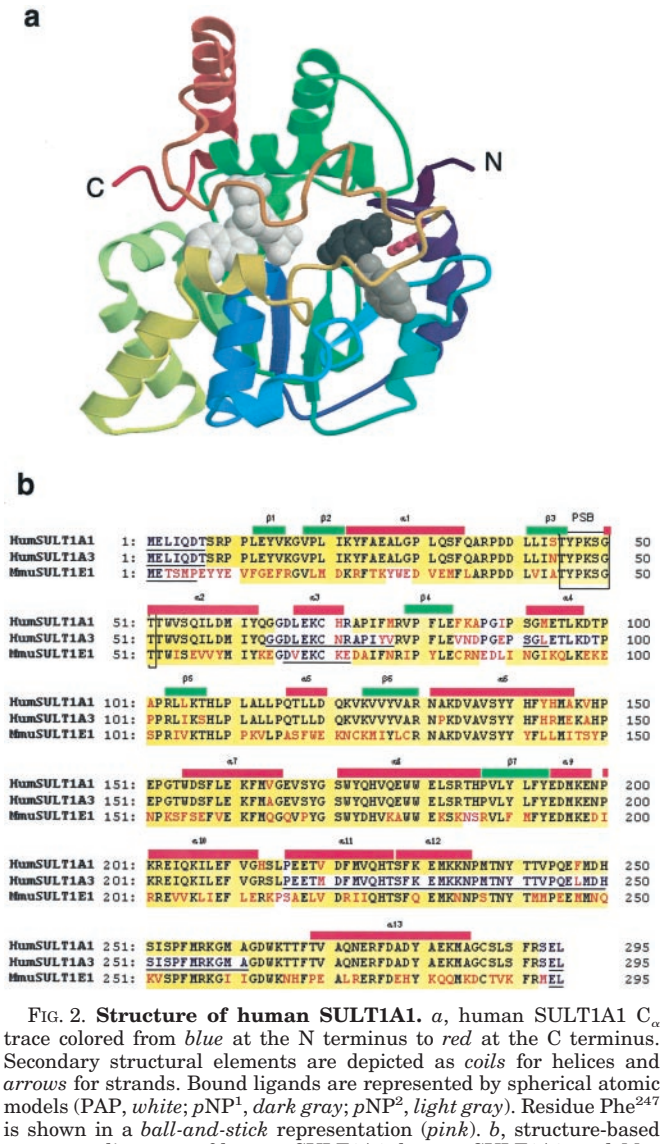


FIG. 2. **Structure of human SULT1A1.** *a*, human SULT1A1 C_α trace colored from *blue* at the N terminus to *red* at the C terminus. Secondary structural elements are depicted as *coils* for helices and *arrows* for strands. Bound ligands are represented by spherical atomic models (PAP, *white*; *p*NP[1], *dark gray*; *p*NP[2], *light gray*). Residue Phe[247] is shown in a *ball-and-stick* representation (*pink*). *b*, structure-based sequence alignment of human SULT1A1, human SULT1A3, and *Mus musculus* SULT1E1. Secondary structural elements in SULT1A1 are indicated by *red* (helix) or *green* (strand) *bars above* the sequence. *PSB* indicates residues in the highly conserved PSB loop. Residues that are not conserved are shown in *red*. Regions of sequences that are structurally similar are *highlighted* with a *yellow background*. Disordered residues in each structure are *underlined*.

structure of the apo form of the highly related SULT1A3 isoenzyme (PDB code 1CJM (10, 16)), which has ~70 disordered residues (23%). A structure-based sequence alignment of the three SULT1 family members for which structures are now known (SULT1A1, SULT1A3, and SULT1E1) is given in Fig. 2*b*.

PAP Binding—The electron density for ligands bound at the active site of SULT1A1 is shown in Fig. 3*a*. The binding mode of PAP to the active site of SULT1A1 is similar to that described previously for the interaction of PAP with mouse SULT1E1 (14). Briefly, 5′-phosphate positioning, thought to be important for orienting the cofactor for in-line sulfuryl transfer to substrate (10, 14, 15), is brought about by an intricate network of interactions with the phosphosulfate binding (PSB) loop (14) of the enzyme ([45]TYPKSGT[51] in SULT1A1). The PAP 3′-phosphate interacts with two conserved regions of sequence (Fig. 2*b*), residues 257-RKG-259 from the G*XX*G*XX*K SULT motif and residues Arg[130] and Ser[138]. As in the SULT1E1 structure, the adenine ring forms stacking and T-shaped interactions with the conserved residues Trp[53] and Phe[229] and is stabilized by hydrogen bond interactions (between N-6 and Thr[227] and between N-3 and Tyr[193]).

Substrate Binding—An unanticipated finding revealed by the crystal structure is the presence of two *p*NP molecules in a large L-shaped substrate binding site in SULT1A1 (Fig. 3*b*). *p*NP[1] (average B-factor, 22 Å^2) is bound in a catalytically competent manner, and its phenol ring is oriented within the active site in the same way that the E2 phenol ring is oriented in the mouse SULT1E1 complex. As in the SULT1E1 structure, Phe[142] and Phe[81] form a substrate access gate that permits binding of planar substrates only at the catalytic site. The phenol hydroxyl of *p*NP[1] forms hydrogen bonds with side chains of catalytic residues His[108] and Lys[106] and with a well ordered water molecule (Fig. 3*b*). The nitro group interacts with a water and forms van der Waals interactions with Val[148], Phe[247], and Met[248]. *p*NP[2] (average B-factor, 41 Å^2) appears to be more weakly bound; no interactions are formed with catalytic residues, but the aromatic ring slots between the side chains of Phe[84] and Phe[76], the nitro group interacts with a water molecule, and van der Waals interactions are formed between the molecule and side chains of Ile[89] and Phe[247].

The L-shaped substrate binding pocket that accommodates both *p*NP[1] and *p*NP[2] is well ordered and very hydrophobic (Fig. 3*b*), incorporating predominantly aromatic residues (Phe[76], Phe[81], Phe[84], Phe[142], His[149], Phe[247], Phe[255], and Tyr[240], Tyr[169], and Phe[24]; the last three residues are not shown in Fig. 3*b* for purposes of clarity) and aliphatic residues (Ile[89], Val[148], Met[248] and, not shown in Fig. 3*b* for clarity, Ile[21], Met[77], Val[243], Pro[90], and Ala[146]). Seven of the 10 aromatic residues of the substrate binding site are conserved in SULT1A3, with the other three replaced either by other aromatic residues or by aliphatic residues (Phe[76] by Tyr, Phe[84] by Val, and Phe[247] by Leu). On the other hand, whereas six of the eight aliphatic residues are conserved as hydrophobic residues in SULT1A3, two of them (Ile[89] and Ala[146]) are replaced by glutamate residues. Residue 146 has already been investigated by site-directed mutagenesis. Thus, a single site mutant of SULT1A1 in which Ala[146] is replaced by glutamate (A146E) shows an ~400-fold reduced affinity for *p*NP (9). Further, a SULT1A3 mutant in which Glu[146] is replaced with the SULT1A1 residue, alanine (E146A), exhibits SULT1A1-like substrate specificity, preferring *p*NP over dopamine (9, 20, 31). These results suggest that the hydrophobic nature of the SULT1A1 binding site is critical for its substrate binding and specificity.

Kinetics of pNP Sulfonation by SULT1A1—The kinetic implications of the presence of two molecules of *p*NP in the active site were investigated using a wide range of concentrations of the substrate (Fig. 4). The results show that there is a slight deviation from Michaelis-Menten kinetics (Fig. 4*a*, *broken line*) at low substrate concentrations, suggestive of some positive cooperativity. However, the most pronounced feature is substrate inhibition at concentrations above 2 μM. These kinetic properties are interpreted under the general model described in Scheme 1. The enzyme can bind *p*NP at site 1 or site 2, and occupancy of site 1 does not prevent subsequent binding at site 2. There are two catalytically competent species, ES_1 and ES_1S_2, and these form product with different efficiencies.

The catalytically competent species ES_1 and ES_1S_2 form enzyme-product complexes (EP and EPS_2) with rate constants of k_1 and k_2, respectively. Release of product from EP is direct (dissociation constant = K_P,) whereas release from EPS_2 requires prior release of *p*NP from site 2 (dissociation constant = K_{PS_2}) (Scheme 1). This equation gives a good fit to the experimental data, as shown by the *solid lines* in Fig. 4. Several

TABLE I
Data measurement and refinement for SULT1A1:PAP:pNP crystal structure
Values in brackets refer to the highest resolution shell (1.90–1.97Å). r.m.s.d., root-mean-square deviation.

Data collection		Refinement	
Unit cell		*R*-factor[b] (%)	18.2
a(Å)	72.8	*R*-free[c] (%)	20.5
b(Å)	124.0	No. of water molecules	154
c(Å)	44.5	Average B-factor (Å^2)	25
α, β, γ (°)	90	Resolution range (Å)	20–1.9
Space group	$P2_12_12$	r.m.s.d. from ideal	
No. of molecules in asymmetric unit	1	Bond length (Å)	0.007
Observations ($I > 0\sigma I$)	82,261	Bond angle (°)	1.3
Unique reflections ($I > 0\sigma I$)	31,483	Ramachandran statistics	
R_{sym} (%)[a]	4.1 (22.0)	Residues in most favored region (%)	92.4
I/σ (I)	8.6 (2.4)	Residues in additionally allowed region (%)	7.6
Completeness (%)	91.4 (63.8%)	Residues in disallowed regions (%)	0
Mosaicity	0.38		

[a] $R_{sym} = \Sigma|I - \langle I\rangle|\Sigma(I)$.
[b] *R*-factor = $\Sigma|F_o - F_c|/\Sigma F_o$.
[c] *R*-free is calculated from 10% of the data.

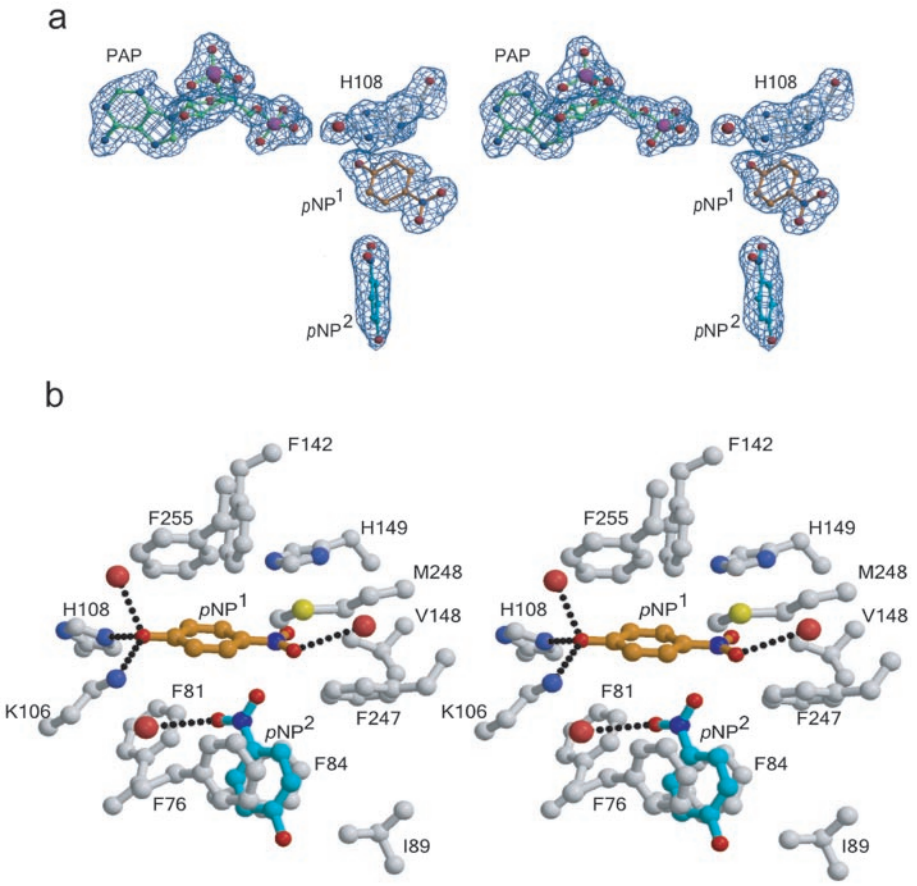


FIG. 3. **SULT1A1 ligand binding.** *a*, stereo view of electron density ($2F_o - F_c$ at 1.1 σ) at the active site of SULT1A1 showing the binding modes of PAP and two molecules of *p*NP. The catalytic residue His^{108} is also shown. *b*, stereo view of the substrate binding site showing the hydrophobic nature of residues surrounding the two *p*NP molecules. Ligands and residues are depicted as *ball-and-stick* models, with atom-based coloring as follows: nitrogen, *dark blue*; oxygen, *red*; phosphorus, *magenta*; sulfur, *yellow*; *p*NP^1 carbon, *orange*; *p*NP^2 carbon, *blue*; PAP carbon, *green*; enzyme carbon, *gray*. Bound water molecules are shown as *red spheres*. Residues not shown for clarity are Ile^{21}, Phe^{24}, Met^{77}, Pro^{90}, Ala^{146}, Tyr^{169}, Tyr^{240}, and Val^{243}.

simpler alternative models were also tested, but these yielded poorer fits to the data. Therefore, we conclude that the presence of two molecules of *p*NP in the active site, as revealed by the structure, represents a real property of SULT1A1 and that the kinetic properties of the enzyme are explained by the influence of occupancy of site 2.

Assuming that all binding steps are at equilibrium and that [P] = 0, the derived rate equation is

$$v = [E]_T[S](k_1K_{s_1s_2} + k_2[S])/(K_{s_1}K_{s_1s_2} + [S]K_{s_1}K_{s_1s_2}(1/K_{s_1} + 1/K_{s_2}) + [S]^2) \quad \text{(Eq. 1)}$$

Equation 1 involves five kinetic constants but can define only four independent parameters (Equation 2) because of internal redundancy.

$$v = V_\infty[S](K_1 + [S])/(K_2K_3 + [S]K_3 + [S]^2) \quad \text{(Eq. 2)}$$

where $V_\infty = k_2[E]_T$, $K_1 = k_1K_{s_1s_2}/k_2$, $K_2 = 1/(1/K_{s_1} + 1/K_{s_2})$, and $K_3 = K_{s_1}K_{s_1s_2}(1/K_{s_1} + 1/K_{s_2})$.

Endogenous Ligands—The size and nature of the binding pocket in SULT1A1 suggests that hydrophobic molecules much larger than *p*NP can be accommodated. Other workers have shown that SULT1A1 catalyzes sulfonation of thyroid hormones such as T2 (11, 12). We therefore undertook docking of T2 into the crystal structure of SULT1A1 and have identified a catalytically competent binding conformation (Fig. 5*a*). In this model, the 4′-OH is oriented toward catalytic residues and the two phenyl rings of T2 adopt an L-shaped conformation with respect to each other, which enables each ring to occupy one of the two *p*NP binding sites. However, this binding mode is only

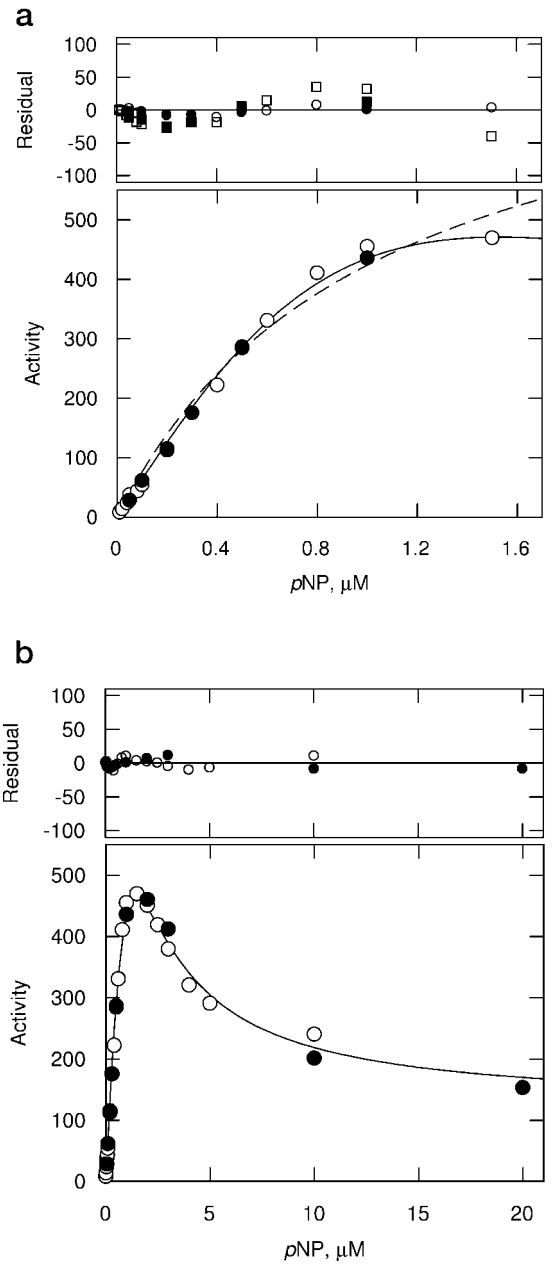
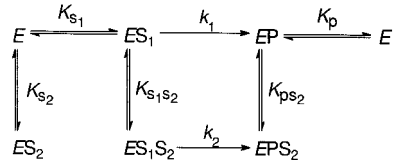


FIG. 4. **Effect of *p*NP concentrations on the activity of SULT1A1.** *Open* and *closed symbols* show the results from two independent experiments. Each data point is the mean of duplicate or triplicate assays, and the standard deviation is contained within the dimensions of the symbols. *Panel a* shows *p*NP concentrations up to 1.5 μM. The *lines* represent a fit to these data of the Michaelis-Menten equation (*broken line*) (V_m = 859 ± 102 nmol min^{-1} mg^{-1}, K_m = 1.0 ± 0.2 μM) or Equation 2 (*solid line*) (V_∞ = 118 ± 18 nmol min^{-1} mg^{-1}, K_1 = 9.0 ± 2.0 μM, K_2 = 16.2 ± 1.4 μM, and K_3 fixed at 0.1 μM). The residuals for each fit are shown above the main graph, with *squares* and *circles* representing the deviations from the Michaelis-Menten equation and Equation 2, respectively. *Panel b* shows the activity of SULT1A1 as a function of *p*NP concentrations up to 20 μM, with the *solid line* defined as described in *a*.

possible if Phe^{247} adopts a different (though still favorable) rotamer conformation to prevent a steric clash with one of the iodine atoms of the substrate (Fig. 5*a*).

SULT1A1 is also reported to sulfonate E2 (3, 13), albeit with lower affinity than other substrates. E2 is a steroid and thus incorporates four fused rings in an extended planar molecular structure. We modeled E2 into the SULT1A1 binding site using the binding mode identified in the SULT1E1-PAP-E2 structure (14). However, E2 cannot be accommodated in this same way in



SCHEME 1. **Kinetic model of SULT1A1 based on the enzyme structure with two molecules of *p*NP in the active site.** The enzyme (E) can bind *p*NP at site 1 to give ES_1 with a dissociation constant of K_{S_1} or at site 2 to give ES_2 with a dissociation constant of K_{S_2}. Occupancy of site 2 prevents *p*NP from binding to site 1, whereas *p*NP at site 1 does not prevent subsequent binding at site 2, to give ES_1S_2.

the crystal structure of SULT1A1 (Fig. 5*b*) because a large number of unfavorable interactions and steric clashes occur between the E2 C/D rings and SULT1A1 even with the altered rotamer conformation for Phe^{247}. We also performed docking studies using GOLD but could not identify a favorable binding mode for E2. The unfavorable interactions that are formed involve residues in two loops (residues 146–154 between α6 and α7 and residues 84–90 just preceding α4; Fig. 5*b*) that close over the hydrophobic active site of SULT1A1 more tightly than they do in the SULT1E1 crystal structure (Fig. 5*c*). This restricts the space available for ligands that bind in an extended conformation in SULT1A1 (Fig. 5*d*) compared with SULT1E1 (Fig. 5*e*).

How is it then that E2, a known substrate of SULT1A1, is accommodated in the binding site? Because the fused ring system of E2 is unlikely to adopt a bent conformation to match the L-shaped binding site in the crystal structure, E2 sulfonation could occur only if there were a conformational change in SULT1A1 that extended the accessible space for multiple fused ring molecules. Thus, the SULT1A1 substrate binding site must be sufficiently plastic to accept both flexible (bent) and rigid (extended) ring systems. The relative K_m values for E2 (85 μM (13)) and T2 (0.5 μM (11)) suggest that binding of the extended fused ring system of E2 by SULT1A1 occurs at an energy cost; this could be due to the conformational rearrangement required for the enzyme to accept the steroid. The changes that would be required to allow E2 binding would most likely involve movement of the two loops described above. We are undertaking co-crystallization trials of SULT1A1-E2 and SULT1A1-T2 to investigate experimentally how the enzyme can accommodate both flexible and rigid multi-ring substrates.

Comparison with SULT1A3 and SULT1E1—At the time of this writing, the crystal structures of three SULT1 family members are known, those of estrogen sulfotransferase (SULT1E1), dopamine sulfotransferase (SULT1A3), and carcinogen-converting sulfotransferase (SULT1A1, reported here). SULTs 1A1 and 1A3 have the highest sequence identity (93%), and yet only 74% of SULT1A1 backbone residues can be superimposed with SULT1A3 (PDB code 1CJM) (212 of the 288 SULT1A1 structurally characterized C_α atoms, root mean square deviation 0.88 Å). By contrast, 93% of residues can be superimposed between SULT1A1 and SULT1E1 (PDB code 1AQU) (269 C_α atoms, root mean square deviation 1.1 Å) although they share only 48% sequence identity. The reason is that 23% of residues in the apoSULT1A3 structure are disordered, whereas very few residues are disordered in the SULT1A1 and SULT1E1 structures (Fig. 6). Both the SULT1A1 and SULT1E1 structures were determined in the presence of cofactor product and substrate. Most of the 68 disordered residues of SULT1A3 correspond to substrate binding regions of SULT1A1 and SULT1E1. These residues adopt very similar conformations in the two ligand-bound enzyme structures (Fig. 6) and make important interactions with both cofactor and substrate ligands (Fig. 3).

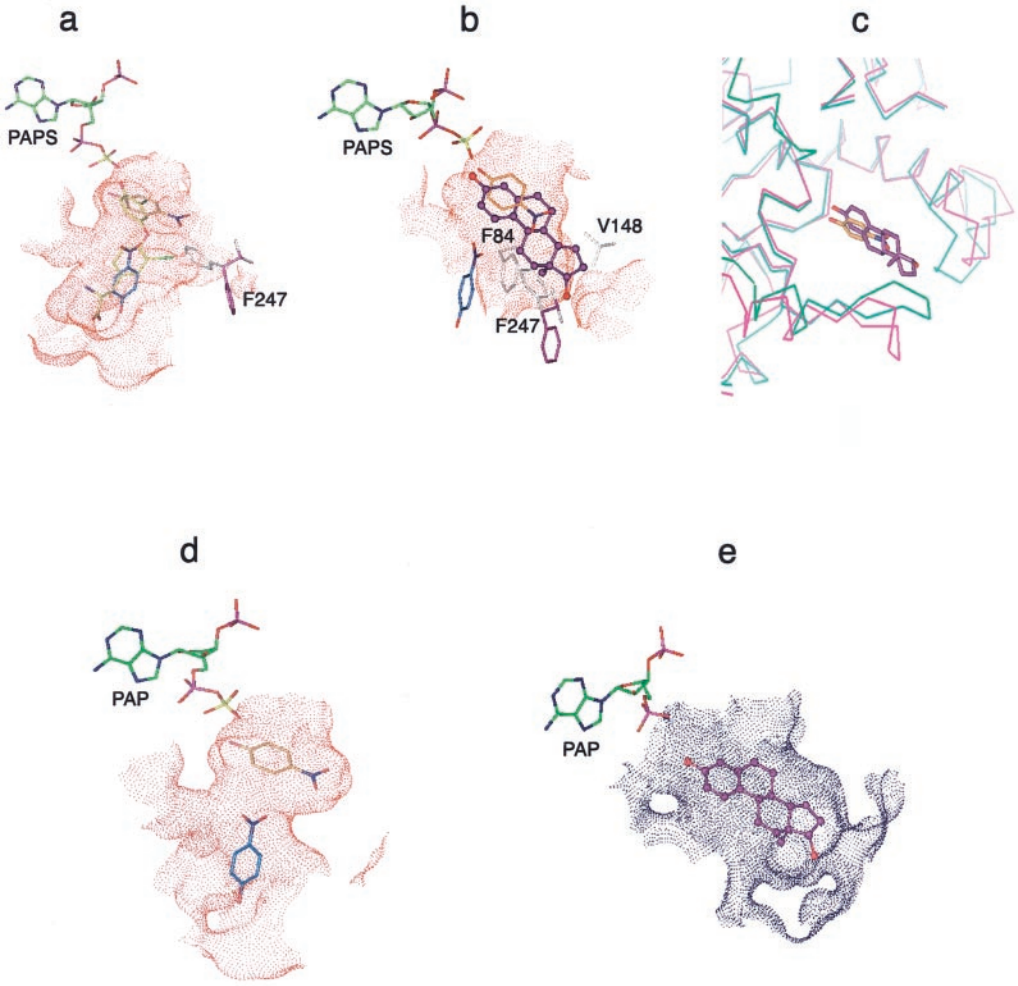


FIG. 5. **E2 and T2 binding.** *a*, T2 (*yellow*) docked into the SULT1A1 active site, showing the solvent-accessible surface calculated with Phe^{247} in an alternate binding mode (*pink*). PAPS is shown in the cofactor binding site. For orientation, the bound conformations of pNP^1 (*orange*) and pNP^2 (*blue*) are also shown. *b*, solvent-accessible surface of SULT1A1 calculated as described in *a*, with E2 (*magenta*) docked into the active site and with PAPS at the cofactor binding site. As in *a*, bound pNP^1 is shown in *orange* and pNP^2 in *blue*. E2 atoms clash with Phe^{247} (in both the x-ray (*gray*) and alternate rotamer conformations (*pink*)), Val^{148} (*gray*), and Phe^{84} (*gray*). *c*, superimposition of the structures of SULT1A1 shown in *green* (with pNP^1 bound, *orange*) and SULT1E1 shown in *pink* (with E2 bound, *magenta*), portraying the striking variation in loop positioning at the active site. *d*, solvent-accessible surface of SULT1A1 at the active site, showing the space available for binding pNP^1 (*orange*) and pNP^2 (*blue*). The position of bound PAP is also shown. *e*, solvent-accessible surface of SULT1E1 at the active site, showing the space available for binding E2 (*magenta*). The position of bound PAP is also shown.

DISCUSSION

Substrate inhibition has been reported previously for both SULT1A1 (32, 33) and SULT1A3 (34) at high concentrations of their preferred small molecule substrates. However, published kinetic studies have generally assumed a Michaelis-Menten model to explain catalysis by these two enzymes and have therefore focused upon moderate substrate concentrations only (35, 36). The crystal structure of SULT1A1 with two molecules of pNP bound at the active site clearly shows how inhibition occurs by this substrate. This is validated by our analysis of SULT1A1 catalysis, which approximates Michaelis-Menten kinetics at low substrate concentrations, characterized by a hyperbolic curve (3, 35, 37, 38) but also exhibits substrate inhibition at high concentrations of pNP. These kinetic properties are consistent with the existence of two binding sites for pNP. Substrate inhibition at high concentrations of pNP is due to impeded catalysis when both binding sites are occupied. The crystal structure of SULT1A1 represents the first sulfotransferase that binds two xenobiotic molecules in the substrate-binding pocket, although an allosteric binding site has previously been postulated for E2 in mouse SULT1E1 (39) and for substrate binding in human thyroid hormone sulfotransferase (SULT1B1) (40). Furthermore, the recently reported structure of DEA-SULT complexed with substrate (19) identifies two alternate substrate binding orientations, one in a catalytic binding mode and the other proposed to give rise to substrate inhibition.

The binding of two molecules of pNP in the active site of SULT1A1 also serves to highlight the large and very hydrophobic nature of the substrate binding region of this enzyme. The 20 residues that surround the two pNP molecules are all aliphatic (Ile, Val, Met, Pro, Ala) or aromatic (Phe, Tyr, His), apart from the catalytic residue Lys-106. It is well known that SULT1A1 can accept a wide assortment of molecules for sulfonation (Fig. 1). These include simple uncharged substituted phenols (9), aromatic hydroxylamines, aromatic heterocyclic hydroxylamines, polycyclic aromatic compounds (8), iodothyronines (11, 12), and estrogens (3, 13). Thus, the substrate binding site can accept small flat aromatic compounds, larger L-shaped aromatics, and extended planar aromatic or aliphatic ring systems. Together, the SULT1A1 crystal structure and our modeling studies show how the substrate binding site can interact with small aromatics and L-shaped aromatics but do not explain easily how extended fused ring systems, such as E2,

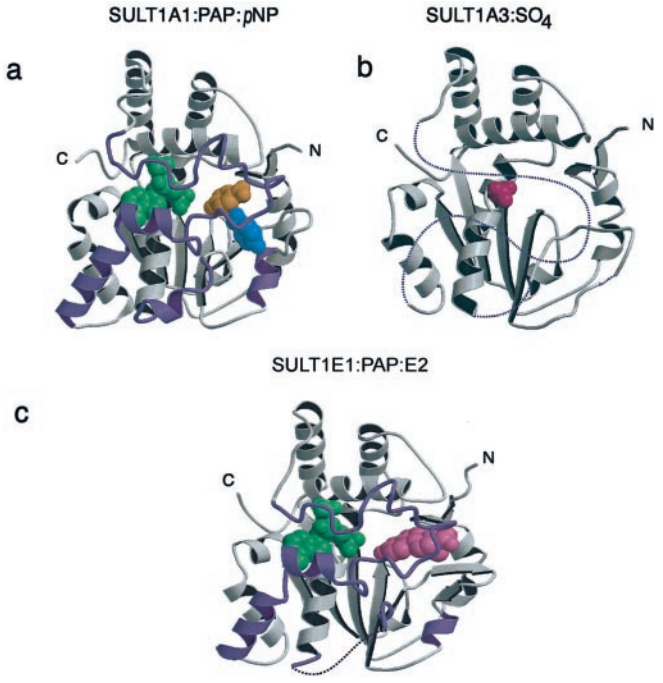


FIG. 6. **Comparison of SULT1 crystal structures.** Structures of human SULT1A1 complexed with PAP (*green*), *p*NP[1] (*orange*), and *p*NP[2] (*blue*) (*a*); human SULT1A3 complexed with sulfate (*pink*) (*b*); and mouse SULT1E1 complexed with PAP (*green*) and E2 (*pink*) (*c*). The *blue* regions in the SULT1A1 structure indicate parts of the protein that are disordered in SULT1A3. The *dotted lines* indicate disordered regions of SULT1A3 and SULT1E1.

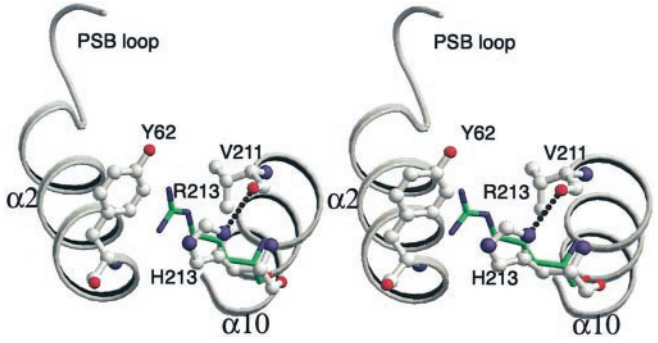


FIG. 7. **Proposed structural basis of SULT1A1 polymorphism.** Stereo view of a region of the SULT1A1 structure near Tyr^{62}, showing the interaction between the His^{213} side chain (found in SULT1A1*2) and the backbone of residue Val^{211}. The more common variant of SULT1A1, SULT1A1*1, has an arginine at position 213, which we propose interacts with Tyr^{62} as found in the structure of SULT1A3 (*green*). The proximity of this region to the PSB loop is shown.

bind to the enzyme. We therefore propose that the substrate binding site of SULT1A1 is plastic, allowing it to adopt varying architectures that will accept a wide range of hydrophobic phenolic compounds including flexible L-shaped molecules and rigid, planar, multiple ring compounds.

Such SULT1A1 active site plasticity may, in part, be a consequence of a disorder-order transition that we propose occurs in the enzyme during catalysis. Thus, the substrate binding regions may be able to adapt to the specific shape of the substrate during the process of binding. Our hypothesis for a disorder-order transition in SULT1A1 is based on the comparison of the crystal structures of the two human enzymes, SULT1A1 and SULT1A3 (Fig. 6). The SULT1A3 structure solved in the absence of cofactor and substrate (10, 16) has a disordered substrate binding site, whereas the SULT1A1-PAP-*p*NP crystal structure complex has an ordered substrate binding pocket. Although the unliganded structure of SULT1A1 has not yet been determined experimentally, the very high sequence conservation between them suggests that the broad structural and mechanistic properties of these two enzymes will also be conserved. Thus, we propose that the apo form of SULT1A1 has a disordered binding site and that during catalysis the enzyme undergoes a disorder-order transition. There are now several examples in the literature of disorder-order transitions occurring during ligand binding or enzymatic catalysis. For example, the type I phosphoribosyltransferase enzymes have a flexible loop that moves to cover the active site during catalysis to protect the transition state from attack by bulk water (41). Similarly, yeast acetohydroxyacid synthase (42, 43), yeast pyruvate decarboxylase (44), ribulose-1,5-bisphosphate carboxylase/oxygenase, and triosephosphate isomerase (45) have been shown to undergo disorder-order transitions at the active site.

Although SULT1A1 and SULT1A3 may share a disorder-order transition because of their high sequence identity, there is sufficient sequence variation between the two to give rise to quite distinct substrate preferences. Thus, although both enzymes accept simple aromatic phenols as substrates, SULT1A3 prefers these substrates to be positively charged, as in dopamine and tyramine (9). In contrast, dopamine and tyramine are relatively poor substrates for SULT1A1, whereas uncharged simple aromatic phenols like *p*NP, *p*-cresol, and *p*-ethylphenol are excellent substrates (9). Furthermore, SULT1A1 is able to accept very large lipophilic substrates such as T2, E2, and polycyclic aromatics, whereas SULT1A3 does not accept these nearly as well. This distinct difference in substrate preference may be due to only two residue differences in the active site. As we have shown, the SULT1A1 active site is extremely hydrophobic, and such an environment would clearly be unfavorable for binding positively charged substrates. However, the size of the SULT1A1 active site is such that it can accept both simple and complex lipophilic substrates (Fig. 1). In SULT1A3, two of the hydrophobic residues of SULT1A1 are replaced with negatively charged residues (Glu^{146} and Glu^{89}), and these would favor binding of substrates with positively charged groups (such as dopamine and tyramine) and discriminate against very hydrophobic compounds, whether large or small. It would be interesting to see whether SULT1A3 would accept a more complex substrate that incorporates a positive charge designed to interact with the acidic residues.

Genetic variation in SULT1A1 has been linked to various cancers because of its ability to sulfonate a wide range of promutagens and procarcinogens. The crystal structure of SULT1A1 reported here is that of variant SULT1A1*2, which has a histidine at residue 213 in place of the more commonly found arginine in SULT1A1*1. The SULT1A1*2 variant is found in about 30% of Caucasians and is reported as having decreased activity and lower stability than SULT1A1*1 (32). This property is thought to contribute to interindividual variation in sulfonation of substrates and hence to the predisposition to cancers. For example, one study (8) investigated the mutagenic activity of two promutagens in a bacterial system expressing either SULT1A1*1 or SULT1A1*2 at roughly the same levels. This study showed that expression of SULT1A1*1 gave an ~10-fold higher rate of mutation than when SULT1A1*2 was expressed. What is the structural effect of replacing Arg^{213} with His? Although we do not have the structure of SULT1A1*1 for comparison, we can investigate this by comparing the structure of SULT1A1*2 with that of SULT1A3, which has an arginine at position 213. Residue 213 is located just before the large disordered/flexible region of SULT1A3 (Fig. 2*b*). In SULT1A1 the histidine at position 213 makes a

hydrogen bond interaction with the main chain carbonyl of residue 211. However, the arginine in SULT1A3 does not form this interaction. Instead it makes a stacking interaction with the side chain of Tyr^{62}, forming what appears to be an amino-aromatic interaction (Fig. 7). Tyr^{62} is highly conserved in the SULT1A family. It would be interesting to mutate Tyr^{62} to a hydrophobic but nonaromatic residue to investigate whether a change in the interacting partner of Arg^{213} would also affect activity in the bacterial mutagenesis system. The position of residue 213, just preceding a flexible region that we propose undergoes disorder-order transition, would explain in part why mutation of this residue could affect enzyme stability. It is also worth noting that Tyr^{62}, the proposed partner for Arg^{213}, is located just before the PSB loop in the enzyme sequence (Fig. 7). This suggests that an alteration in the Tyr^{62}–Arg^{213} interaction could impact on cofactor binding.

In summary, the structure of SULT1A1 reveals an extended and very hydrophobic binding site, which is consistent with binding of both small and large phenolic substrates and explains the observed substrate inhibition with *p*NP. Binding of extended rigid substrates like E2 may require active site plasticity and a disorder-order transition, which we propose is conserved between SULT1A1 and SULT1A3. Furthermore, the structure of SULT1A1 will now help us to unravel how genetic variation in this enzyme translates into interindividual variation in the predisposition cancer.

Acknowledgment—We thank Joel Tyndall for expert assistance with GOLD.

REFERENCES

1. Debaun, J. R., Miller, E. C., and Miller, J. A (1970) *Cancer Res.* **30,** 115–127
2. Glatt, H. (1997) *FASEB J.* **11,** 314–321
3. Falany, C. N. (1997) *FASEB J.* **11,** 1–2
4. Wang, Y., Spitz, M. R., Tsou, A. M., Zhang, K., Makan, N., and Wu, X. (2002) *Lung Cancer* **35,** 137–142
5. Bamber, D. E., Fryer, A. A., Strange, R. C., Elder, J. B., Deakin, M., Rajagopal, R., Fawole, A., Gilissen, R. A., Campbell, F. C and Coughtrie, M. W. (2001) *Pharmocogenetics* **11,** 679–685
6. Seth, P., Lunetta, K. L., Bell D. W., Gray, H., Nasser, S. M., Rhei, E., Kaelin, C. M, Iglehart, D. J., Marks, J. R., Garber, J. E, Haber, D. A., and Polyak, K. (2000) *Cancer Res.* **60,** 6859–6863
7. Weinshilboum, R. M., Otterness, D. M., Aksoy, I. A., Wood, T. C, Her, C., and Raftogianis, R. B. (1997) *FASEB J.* **11,** 3–14
8. Glatt, H., Boeing, H., Engelke, C. E., Ma, L., Kuhlow, A., Pabel, U., Pomplun, D., Teubner, W., Meinl, W. (2001) *Mutat. Res.* **482,** 27–40
9. Brix, L. A., Barnett, A. C., Duggleby, R. G., Leggett, B., and McManus, M. E. (1999) *Biochemistry* **38,** 10474–10479
10. 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
11. Li, X., Clemens, D. L., Cole, J. R., and Anderson, R. J. (2001) *J. Endocrinol.* **171,** 525–532
12. Stanley, E. L., Hume, R., Visser, T. J., and Coughtrie, M. W. (2001) *J. Clin. Endocrinol. Metab.* **86,** 5944–5955
13. Adjei, A. A., Weinshilboum, R. M. (2002) *Biochem. Biophys. Res. Commun.* **292,** 402–408
14. Kakuta, Y., Pedersen, L. G., Carter, C. W., Negishi, M., and Pedersen, L. C. (1997) *Nat. Struct. Biol.* **4,** 904–908
15. Pedersen, L. C., Petrotchenko, E., Shevtsov, S., and Negishi, M. (2002) *J. Biol. Chem.* **277,** 7928–7932
16. 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
17. Pedersen, L. C., Petrotchenko, E. V., and Negishi, M. (2000) *FEBS Lett.* **475,** 61–64
18. Kakuta, Y., Sueyoshi, T., Negishi, M., and Pedersen, L. C. (1999) *J. Biol. Chem.* **274,** 10673–10676
19. Rehse, P. H., Zhou, M., and Lin, S.-X. (2002) *Biochem. J.* **364,** 165–171
20. Liu, M. C., Suiko, M., and Sakakibara, Y. (2000) *J. Biol. Chem.* **275,** 13460–13464
21. Gaedigk, A., Lekas, P., Berchuck, M and Grant, D. M. (1998) *Chem. Biol. Interact.* **109,** 43–52
22. Foldes, A., and Meek, J. L (1973) *Biochim. Biophys. Acta* **327,** 365–374
23. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54,** 905–921
24. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjelgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
25. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–391
26. Hooft, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature* **381,** 272
27. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig H., Shindyaloy, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* **28,** 235–242
28. Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997) *J. Mol. Biol.* **267,** 727–748
29. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 940–950
30. Merritt, E. A., and Bacon, D. (1997) *Methods Enzymol.* **277,** 505–524
31. Dajani, R., Hood, A. M., and Coughtrie, M. W. (1998) *Mol. Pharmacol.* **54,** 942–948
32. Raftogianis, R. B., Wood, T. C., and Weinshilboum, R. M. (1999) *Biochem. Pharmacol.* **58,** 605–616
33. Reiter, C., Mwaluko, G., Dunnette, J., Van Loon, J., and Weinshilboum, R. (1983) *Naunyn-Schmiedeberg's Arch. Pharmacol.* **324,** 140–147
34. Ganguly, T. C., Krasnykh, V., and Falany, C. N. (1995) *Drug Metab. Dispos.* **23,** 945–950
35. Brix, L. A., Duggleby, R. G., Gaedigk, A., and McManus, M. E. (1999) *Biochem. J.* **337,** 337–343
36. Lewis, A. J., Kelly, M. M., Walle, U. K., Eaton, E. A., Falany, C. N., and Walle, T. (1996) *Drug Metab. Dispos.* **24,** 1180–1185
37. Honma, W., Kamiyama, Y., Yoshinari, K., Sasano, H., Shimada, M., Nagata, K., and Yamazoe, Y. (2001) *Drug Metab. Dispos.* **29,** 274–281
38. Veronese, M. E., Burgess, W., Zhu, X., and McManus, M. E. (1994)) *Biochem. J.* **302,** 497–502
39. Zhang, H., Varlamova, O., Vargas, F. M., Falany, C. N., Leyh, T. S., and Varmalova, O. (1998) *J. Biol. Chem.* **273,** 10888–10892
40. Wang, J., Falany, J. L., and Falany, C. N. (1998) *Mol. Pharmacol.* **53,** 274–282
41. Schumacher, M. A., Carter, D., Ross, D. S., Ullman, B., and Brennan, R. G. (1996) *Nat Struct. Biol.* **3,** 881–887
42. Pang, S. S., Duggleby, R. G., and Guddat, L. W. (2002) *J. Mol. Biol.* **317,** 249–262
43. Pang, S. S., Guddat, L. W., and Duggleby, R. G. (December 20, 2002) *J. Biol. Chem.* 10.1074/jbc.M211648200
44. Lu, G., Dobritzsch, D., Baumann, S., Schneider, G., and Konig, S. (2000) *Eur. J. Biochem.* **267,** 861–868
45. Schreuder, H. A., Knight, S., Curmi, P. M., Andersson, I., Cascio, D., Branden, C. I., and Eisenberg, D. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 9968–9972