High Accuracy in Silico Sulfotransferase Models*§

Received for publication, August 21, 2013, and in revised form, October 6, 2013. Published, JBC Papers in Press, October 15, 2013, DOI 10.1074/jbc.M113.510974

Ian Cook‡, Ting Wang‡, Charles N. Falany§, and Thomas S. Leyh††

From the ‡Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461-1926 and the §Department of Pharmacology and Toxicology, University of Alabama School of Medicine at Birmingham, Birmingham, Alabama 35294-0019

Background: Human sulfotransferases (SULTs) regulate the bioactivities of hundreds of compounds in vivo.

Results: The in silico models of SULTs developed here proved remarkably accurate in identifying SULT substrates in a 1,455-compound library containing all FDA-approved drugs.

Conclusion: Highly accurate in silico SULT models are now available.

Significance: The elision of mechanism and modeling can produce remarkably accurate in silico tools for the study of biology.

Predicting enzymatic behavior in silico is an integral part of our efforts to understand biology. Hundreds of millions of compounds lie in targeted in silico libraries waiting for their metabolic potential to be discovered. In silico “enzymes” capable of accurately determining whether compounds can inhibit or react is often the missing piece in this endeavor. This problem has now been solved for the cytosolic sulfotransferases (SULTs). SULTs regulate the bioactivities of thousands of compounds—endogenous metabolites, drugs and other xenobiotics—by transferring the sulfuryl moiety (SO3) from 3′-phosphoadenosine 5′-phosphosulfate to the hydroxyls and primary amines of these acceptors. SULT1A1 and 2A1 catalyze the majority of sulfation that occurs during human Phase II metabolism. Here, recent insights into the structure and dynamics of SULT binding and reactivity are incorporated into in silico models of 1A1 and 2A1 that are used to identify substrates and inhibitors in a structurally diverse set of 1,455 high value compounds: the FDA-approved small molecule drugs. The SULT1A1 models predict 76 substrates. Of these, 53 were known substrates. Of the remaining 23, 21 were tested, and all were sulfated. The SULT2A1 models predict 22 substrates, 14 of which are known substrates. Of the remaining 8, 4 were tested, and all are substrates. The models proved to be 100% accurate in identifying substrates and made no false predictions at Kd thresholds of 100 µM. In total, 23 "new" drug substrates were identified, and new linkages to drug inhibitors are predicted. It now appears to be possible to accurately predict Phase II sulfonation in silico.

The Phase I and II metabolizing enzymes provide a critical first line of defense against the chemical insults of xenobiotics (1). Considerable effort has gone into the development and testing of in silico models that predict the binding and catalytic properties of these enzymes (2). The cytochrome P-450 isozymes that catalyze the oxidative reactions of Phase I metabolism are of particular interest because they are estimated to metabolize ~85% of drugs (1). Sulfotransferases (SULTs) and UDP-glucuronosyltransferases catalyze the majority of the Phase II conjugation reactions, and together they conjugate ~40% of drugs (1). Attempts to model glucuronosyltransferases have been described (3); however, to our knowledge, no attempt to develop in silico models that predict the sulfation component of Phase II metabolism has been reported.

The human cytosolic SULTs comprise a small (<13-member) enzyme family that catalyzes transfer of the sulfuryl moiety (SO3) from 3′-phosphoadenosine 5′-phosphosulfate (PAPS) to the hydroxyls and primary amines of thousands of recipients: metabolites, drugs, and other xenobiotics (4). The in vivo activities of these compounds are regulated by sulfation, which often profoundly alters their target affinities (5–8). SULTs perform at least two essential metabolic functions: a homeostatic function, in which, for example, they regulate the receptor binding activities of peptide and steroid hormones (6, 7), and a defensive function (9, 10), in which they sulfonate the myriad compounds that pass through the liver and would otherwise adventitiously bind receptors and regulate cellular signaling systems.

Predicting metabolism is a major objective of biological research, and in silico prediction of the molecular behavior of enzymes is an integral part of this effort (2). Here, experimental data and recent insights into the molecular basis of SULT substrate selectivity were used to develop and benchmark in silico models that can predict the binding and reactivity of two SULTs: 1A1 and 2A1. These two SULT isoforms are each present in near gram quantities in a typical adult liver (where they comprise ~80% of SULTs by mass) (11) and are responsible for the majority of the sulfation that occurs during first pass metabolism. The accuracy of the models was tested by using them to predict the binding and reactivity of a moderately large set of structurally diverse compounds—the 1,455 FDA-approved drugs (12)—and then testing the predictions experimentally. Both models proved to be 100% accurate in identifying substrates, and neither made false predictions using a ligand affin-

* This work was supported, in whole or in part, by National Institutes of Health Grants GM54469 and GM38953.

This article contains supplemental Tables S1 and S2 and supplemental references.

†1 To whom correspondence should be addressed: Dept. of Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461-1926. Tel.: 718-430-2857; Fax: 718-430-8711; E-mail: tsleyh@gmail.com.
ity threshold of 100 μM. Together, the models identified 98 SULT substrates in the drug library, 23 of which are identified here for the first time. These hyperaccurate models are expected to provide valuable tools for the in silico exploration of sulfur metabolism.

**EXPERIMENTAL PROCEDURES**

**Materials**

The materials and their sources are as follows. 4-Nitrophenol (PnP), DTT, EDTA, l-glutathione (reduced, GSH), glucose, imidazole, isopropyl thio-β-ᴅ-galactopyranoside, LB media, lysozyme, β-mercaptoethanol, pepstatin A, and potassium phosphate were the highest grade available from Sigma. Silica Gel (60 Å) and PEI-F TLC plates (Whatman and EMD), ampicillin, HEPES, KOH, MgCl₂, NaCl, KCl, and phenylmethylsulfonyl fluoride were purchased from Fisher Scientific. Glutathione- and nickel-chelating resins were obtained from GE Healthcare. Competent _Escherichia coli_ (BL21(DE3)) was purchased from Novagen. FDA-approved drugs were purchased from Enzo Life Science. PAP and PAPS were synthesized as previously described (13).

**Software and Computational Equipment**

Simulations were performed on a Parallel Quantum Solutions QS32–2670C-XS8 computer. MODELLER was provided by the University of California, San Francisco. A GOLD license was obtained from the Cambridge Crystallographic Data Center. The Approved Drug 3.0 in silico library was obtained from the DrugBank database at the University of Alberta (Edmonton, Canada) (12). The source code for GROMACS 4.5 was downloaded from the GROMCAS General Public License. AMBER and Ambertools 10.0 were obtained from the University of California, San Francisco.

**Methods**

**Protein Purification**—Human SULT DNA was codon optimized for _E. coli_ (Mr. Gene, Regensburg, Germany) and inserted into a triple-tag pGEX-6P expression vector with an N-terminal His/GST/maltose-binding protein tag (13, 14). The plasmid was transfected into _E. coli_ (BL-21(DE3)), expressed, and purified as described previously (15). Briefly, the cell pellet was suspended in lysis buffer, sonicated, and centrifuged. The supernatant was loaded onto a chelating Sepharose Fast Flow column charged with Ni²⁺. The fusion protein was eluted with imidazole (10 mM) onto a glutathione-Sepharose column and then eluted using GSH (10 mM). The fusion protein was digested with Precision Protease and diazylated overnight against HEPES/K⁺ (50 mM, pH 7.5), DTT (1.5 mM), KCl (50 mM) at 4 °C. The sample was passed back through the glutathione column to remove the tag. SULT1A1 was concentrated using a 10-kDa cutoff filter and stored at −20 °C in 40% glycerol. Protein purity was assayed at >97% using SDS-PAGE. Protein concentration was determined spectrophotometrically (ε₂₈₀ = 36.7 mm⁻¹ cm⁻¹) (16).

**Molecular Dynamic Simulations**—Models of SULT2A1 with bound nucleotide (PAP) were constructed from the available crystal structures (4GRA for 1A1 and 1EFH for 2A1) (17, 18). Missing atoms were added using MODELER (19). For simulations involving PAPS-bound enzyme, the PAPS structure was obtained from Protein Data Bank entry 1HY3 (20), its charge distribution was calculated using AmberTools 10.0, and PAP was replaced with PAPS (21). The protein was protonated and solvated using GROMACS (22). The charge was neutralized with Na⁺, and NaCl was added to the solvent for a concentration of 0.15 M. “Steepest descents” in GROMACS was used to energy minimize the system (23). Once minimized, the protein-solvent system was heated to a simulated temperature of 310 K, and the system was stabilized using Berendsen temperature and pressure coupling (24). Bonds were constrained with LINCS (25). The simulation was equilibrated and run for 10 ns. A model of the enzyme was then generated from the simulations using the _g_cluster_ function in GROMACS (26). This was done for SULT1A1 and 2A1 with and without PAPS.

**The DrugBank Library**—The screens described in this study used the DrugBank Approved Drug Library 3.0. This library contains the 1,455 small molecules (≥1,000 Da) that are approved for therapeutic use in any country and includes the complete list of FDA-approved small molecule over the counter (390 drugs) and prescription (821 drugs) drugs listed in the 31st edition of the Orange Book (27).

**Literature Search Strategy for Prior Sulfation Studies on DrugBank Compounds**—To identify prior sulfation studies performed on the compounds listed in the DrugBank library, the PubMed and DrugBank (12, 27) databases and the FDA drug labels were searched using the following search function: (drug name given in the DrugBank library) and (sulfation or sulfation of SULT).

**Substrate Screening**—Compounds in the DrugBank list that had no literature link to sulfation and were predicted to be substrates by the SULT1A1 and 2A1 models were purchased from the Enzo drug library (28) and tested as substrates under following conditions: SULT (0.10 μM dimer), substrate (10 μM), MgSO₄ (0.1%, 35S-PAPS (3.0 μM, 0.69 Ci/mmol), MgCl₂ (5.0 mM), KPO₄ (50 mM), pH 7.5, and T = 25 ± 2 °C. The reactions were initiated by the addition of PAPS. The final reaction volume was 30 μl. The reactions were quenched after 16 h with 0.10 M NaOH and then neutralized with HCl. Tubes containing the samples were placed in a boiling water bath for 1 min before centrifuging at 12,100 relative centrifugal force for 5 min. The samples were spotted onto a reverse phase TLC plate and reactants were separated using the following running buffer: methylene chloride, methanol, water, and ammonium hydroxide (90, 16, 3.5 and 0.50% by volume, respectively). Radiolabeled products were visualized and quantitated using a STORM imaging system.

**Thioguanine Binding and Inhibition**—To assess whether thioguanine binds and inhibits SULT1A1, it was tested as an inhibitor of the SULT1A1 catalyzed sulfation of PnP. Sulfation was monitored at 415 nm (ε = 10.6 mm⁻¹ cm⁻¹, pH 7.5 (29)). The reaction conditions were as follows: SULT1A1 (50 mM), PnP (1.5 μM, 1 × K⁺ₘ₀), inhibitor (0, 50, 100, 200, 500, 1,000, and 1,500 μM), PAPS (100 μM), MgCl₂ (5.0 mM), KPO₄ (50 mM), pH 7.5, 25 ± 2 °C. Reactions were initiated by addition of PAPS. Reaction rates were determined during conversion of the first
The Sulfation of FDA-approved Drugs

5% of PnP. $K_i$ values were obtained by fitting the rate versus [inhibitor] data to a single-site binding model (15).

Thioguanine Reactivity—Reactions were initiated by mixing a solution containing SULT1A1 (30 μM, 83.5 × $K_m$ PAPS), MgCl₂ (5.0 mM), KPO₄ (50 mM), pH = 7.5 with an equal volume of a solution containing thioguanine, ³⁵S-PAPS (3.0 μM, SA 0.69 Ci/mmol), MgCl₂ (5.0 mM), KPO₄ (50 mM), pH = 7.5. The concentration of thioguanine was selected, based on its $K_i$ value to saturate the enzyme. Reactions were run for 2.0 min, and quenched by addition of NaOH to a final concentration of 0.10 M and then neutralized with HCl. Eppendorf tubes containing the reaction mixtures were heated in a boiling water bath for 1 min and then centrifuged for 5 min at 12,000 relative centrifugal force. Radiolabeled reactants were separated on anion exchange TLC plates using a 0.90M LiCl mobile phase and quantitated using STORM imaging. Controls were identical except that acceptor was not added.

Amoxapine and Protriptyline Inhibition of SULT1A1—Inhibition was assessed in a classical initial rate study using a progress curve strategy (30). In this approach, the inhibitor concentration is fixed, and the concentration of the substrate is continuously monitored from $t_0$ to the end point of the reaction. Progress curves are collected at four inhibitor concentrations that range from 0.2 – 5 × $K_i$. The velocity at each point in the curve is given by the slope of the tangent at that point. Slopes taken over a sufficiently small concentration range (∼5% of remaining substrate) provide initial rates. Thousands of initial rate measurements can be calculated from a properly executed and analyzed progress curve, and the data can be plotted and analyzed in the familiar Lineweaver-Burke, double-reciprocal format.

Reaction conditions were as follows: SULT1A1 or 2A1 (50 nM), PnP (3.0 or 100 μM, respectively; 2 × $K_m$ PnP), amoxapine or protriptyline (0, 50, 100, or 300 μM), PAPS (125 μM or 3.0 mM, respectively; ≥250 × $K_m$ PAPS), MgCl₂ (5.0 mM), KPO₄ (50 mM), pH 7.5, 25 ± 2 °C. Sulfation of PnP was monitored optically at λ = 415 nm ($ε_{415} = 10.6 \text{ mm}^{-1} \text{ cm}^{-1}$ at pH 7.5 (29)). The affinities of PAPS and PAP for 1A1 are comparable; hence, product inhibition by nucleotide was minimized by selecting a PnP concentration that converts ∼24% of PAPS to PAP at the reaction endpoints. At the end point, the concentration of PnP sulfate is 400-fold lower than its $K_i$ (1.2 mM (31)); hence, PNP sulfate inhibition should also be negligible. The progress curve was analyzed using methods described in previous work (30).

Briefly, an initial rate is calculated at each point in the progress curve by taking the slope over an interval that is centered on the point and represents 1% of the remaining absorbance. The velocities and concentrations are plotted in double reciprocal space and fit globally to the competitive binding algebra using the Cleland statistical analysis algorithm (32) modified to accept large data sets (30).

RESULTS AND DISCUSSION

Selecting Docking Scaffolds—SULTs contain a conserved ~30-residue active site that restructures in response to substrates and mediates their interactions. The cap plays a pivotal role in substrate selection. In the open state, the cap allows active site access to a broad range of acceptor structures. When it closes, it forms a pore that sterically sieves acceptors from its environment, admitting only acceptors small enough to pass through and the active site. Greater than 95% of the ligand-free enzyme is in the cap-open form (15, 33). In contrast, the binding of nucleotide causes the majority of the enzyme to shift into the closed state; the isomerization equilibrium constant, $K_{iso}$, for the nucleotide-bound complex is ~24 in favor of the closed form. A consequence of this isomerization is that the affinities of substrates that are too large to pass through the pore are weakened by bound nucleotide by a factor that is nearly equal to $K_{iso}$ (i.e., $K_d(+:nuc) = (1 + K_{iso}) \cdot K_d(-:nuc)$ (15)). In liver cytosol, where PAPS concentration (~80 μM (34)) greatly exceeds its affinity constant (~0.3 μM (15)), SULTs 1A1 and 2A1 are expected to be largely in the closed form. The dimensions of the pore define whether substrates are “large” or “small,” and the large/small bias of the enzyme selection is determined by the value of $K_{iso}$. Given these facts, it is clear that comprehensive modeling of SULT binding and reactivity requires inclusion of both the open and closed forms.

Docking—Each of the FDA-approved small molecule drugs in the DrugBank database was docked into rigid structures of the open and closed forms of SULT1A1 and 2A1 using GOLD, the Genetically Optimized Ligand Docking program (35). The protein structures used in the docking were obtained by analyzing the ensemble of structures predicted to occur at thermal equilibrium using GROMACS. The ensemble was binned into clusters whose all protein atom root mean square deviations vary by no more than 2 Å. The centroid of the largest cluster in an ensemble, the structure that differs least from all other members in the set, was used in the docking experiments. Centroids of the closed forms were generated from crystal structures of nucleotide-bound enzyme. Open form centroids were derived from their closed form counterparts by removing the nucleotide and re-equilibrating (see “Experimental Procedures”). This procedure is known to convert the closed into the open form (17). In the case of SULT2A1, the open centroid closely resembles the structure of the unliganded enzyme (17). There are no other unliganded SULT structures. GOLD morphs ligands in the active site cavity by swapping bond angle and position parameters of two randomly configured versions of the same molecule located within a sphere centered on the active site. Swapping occurs with a preset “cross-over” frequency. Changes that are not possible via swapping are allowed to occur with a preset “mutational” frequency. Parents and siblings from a given “cross” are scored on the basis of stearics and energetics, and the top scoring half is used in the subsequent cross. This cycle is repeated 2,500 times or until the program converges on a single, most stable ligand configuration. The convergence routine is repeated 10 times for each ligand-enzyme combination, and the most stable result is used for subsequent analysis.

Selecting the Binding Cutoff—The discovery that a compound is a SULT substrate will likely be physiologically relevant only if the $K_i$ or $K_m$ of the substrate is comparable to or less than its $K_{iso}$.

The in vivo concentration is the concentration (or activity) of free ligand in the near environment of an enzyme in a particular local in the organism; for example, in the hepatocyte cytosol. In practice, these values are rarely known. $K_m$ is often taken as an approximation of in vivo sub-
Experimental and calculated binding free energy correlations. A, SULT1A1 correlation. B, SULT2A1 correlation. Experimental binding free energy (BFE) values were calculated using ligand affinity constants determined in fluorescence titrations that monitor binding via ligand induced changes in intrinsic protein fluorescence. The ligands used in constructing the plots in A and B and their respective calculated and experimentally determined binding free energies (kcal/mol) are as follows: A, apomorphine, –13, –19 (8); acetaminophen, –12, –17 (37); ethyl estradiol, –12, –15 (37); 2-naphthol, –11, –17 (37); 4-nitrophenol, –11, –13 (17); N-propyl apomorphine, –11, –15; 17β-estradiol, –10, –13 (8); fulvestrant, –7.9, –12 (17); raloxifene, –7.5, –11 (38); B, 24(5)-OH-droloxysterol, –10.6, –8 (5); 22(24)-OH-hyroxysterol, –10.3, –7.7 (5); 17β-estradiol, –9.7, –6.5 (40); emodin, –9.5, –7.1 (41); ethyl estradiol, –9.5, –6.8 (37); androstenedione, –9.0, –6.8 (42); DHEA, –8.6, –6.9 (15); pregnenolone, –7.9, –7.2 (37); raloxifene, –7.2, –6.6 (15); lithocholic acid, –5.8, –6.1 (43); fulvestrant, –2.2, –4.3 (44); fentolol, –1.3, –4.1 (45); 22(5)-hydroxycholesterol, –0.9, –4.0 (5).

Distinguishing Inhibitors from Substrates—Inhibitors were distinguished from substrates on the basis of whether atoms that might attack the sulfuryl moiety are properly positioned in the catalytic machinery to accomplish the transfer chemistry. Cytosolic SULTs contain a universally conserved active site histidine (His-99 and His-106 in SULTs 1A1 and 2A1, respectively) that activates attack by abstracting a proton from the nucleophilic hydroxyl or primary amine of the acceptor. If the nucleophilic atom (nitrogen or oxygen) is ≥4 Å from either the δ- or ε-N of the catalytic His, it was considered unlikely to engage in chemistry, and the compound was categorized as an inhibitor; if it fell within the cutoff, the compound was scored as a substrate. Compounds were docked 10 times and scored as a substrate if the nucleophilic atom fell within the 4-Å cutoff in any docking outcome.

Predicted 1A1 Substrates—Of the 1,455 FDA-approved small molecule drugs contained in the DrugBank library, 76 were predicted to be SULT1A1 substrates. Of the 76 compounds, 53 were reported substrates for SULT1A1, and the remaining 23 had not been identified. Of the 23, 21 were purchased and tested as substrates. The remaining two were prohibitively expensive. Reaction conditions were as follows: SULT1A1 (0.10 μM), putative substrate (10 μM), 35PAPS (3.0 μM), MgCl2 (5.0 mM), KPO4 (50 mM), pH 7.5, 25 ± 2 °C. Reactions were quenched and spotted onto reverse phase TLC plates, and the radiolabeled reactants were separated and then quantitated using a STORM imaging system. Of the 21 test compounds, only thioguanine (lane 22) did not yield a band of sulfonated product (Fig. 2).

To determine whether thioguanine bound to SULT1A1, it was tested in a classical initial rate study as an inhibitor of PnP sulfation (see “Experimental Procedures”). Thioguanine inhibited PnP sulfation competitively with a Ki of 70 ± 10 μM. The pKi of thioguanine is low, 3.7 (46). Sulfonated, low pKi, primary amines are typically unstable toward hydrolysis (47). If the sulfonated product were hydrolyzed quickly relative to its formation, product will not be observed; instead, thioguanine will appear to stimulate production of 35SO4 from 35PAPS. This is precisely what is observed. Under the conditions described in the Fig. 3 legend, the addition of thioguanine stimulates SO4 formation 30-fold over background.

In summary, of the 76 predicted substrates, 74 were tested, either in the current work or in previous studies. All 74 were substrates. The 21 drugs that were not previously identified as substrates are listed in Table 1. A complete listing of the 76 substrates, along with their predicted binding free energies, a literature reference (where available), and the form of the enzyme to which they bind, can be found in supplemental Table S1. The algorithms made no false positive predictions (all predicted substrates were indeed substrates) and no false negative predictions (our literature search (see “Experimental Procedures”) did not identify any drugs in the database that were sulfated beyond those predicted to be substrates).

Predicted 2A1 Substrates—The in silico screen predicted 22 SULT2A1 substrates, 8 of which have not been reported in the literature (supplemental Table S1). Of the predicted compounds, 4 were prohibitively expensive (~$3,000/mg). The remaining 4 were tested using the following reaction condi-
The Sulfation of FDA-approved Drugs

Reactions were quenched, and the radiolabeled reactants were separated using reverse phase TLC and quantitated using STORM imaging. All four drugs proved to be substrates (Fig. 4 and Table 1). Dienestrol (Fig. 4, lane 4) and arbutamine (lane 5) contain multiple hydroxyls, and the complex TLC banding patterns observed with these compounds strongly suggest that they are sulfated at multiple sites. These two compounds are substrates for both SULT1A1 and 2A1, and it is interesting to note that the SULT1A1 product(s) migrates as a single band using the identical TLC system (Fig. 2, lanes 11 and 16). The differences in the complexity of the 1A1 and 2A1 products underscore the different substrate specificities of these isoforms. Of the 22 predicted substrates, 18 have been tested, and all are confirmed substrates. Here again, the algorithm was 100% accurate and gave no false positive or false negative predictions. The 22 predicted substrates are listed in supplemental Table S1, along with their calculated binding free energies, a literature reference (where appropriate), and the enzyme form(s) to which the compound binds.

The Inhibition of SULTs—Phase II metabolism occurs in a complex environment in which hundreds of compounds (inhibitors, metabolites, drugs, nutrients, and other xenobiotics) compete for the active sites of a handful of SULT isozymes. In certain cases, SULT inhibition is the molecular basis of drug-drug interactions (48, 49), which can have profound clinical consequence (50). In the hope of expanding our understanding of this important issue, the open and closed models of SULT1A1 and 2A1 were used to screen the drug library for possible inhibitors.

**FIGURE 2. Testing predicted SULT1A1 substrates.** Twenty-two FDA-approved drugs not previously identified as SULT substrates were tested as SULT1A1 substrates. The figure shows the STORM image of a TLC sheet spotted with reactions in which 35PAPS was used to label the predicted substrates. The lane loading scheme is as follows: lane 1, negative control (— acceptor); lane 2, betazole; lane 3, naloxone; lane 4, fenoterol; lane 5, naltrexone; lane 7, levonorethodrine; lane 8, enoxacin; lane 9, pyrimethamine; lane 10, aminoglutethimide; lane 11, arbutamine; lane 12, edrophonium; lane 13, monobenzone; lane 14, betaniol; lane 15, mimosine; lane 16, dienestrol; lane 17, levallorphan; lane 18, naltrexone; lane 19, desvenlafaxine; lane 20, furosemide; lane 21, clonidine; lane 22, thioguanine; lane 23, estradiol (positive control). The reaction conditions were: SULT1A1 (0.10 μM), putative substrate (0 or 10 μM), 35PAPS (3.0 μM), MgCl2 (5.0 mM), KPO4 (50 mM), pH 7.5, 25 °C. The reactions were quenched and spotted onto reverse phase TLC plates, and reactants were separated and quantitated using STORM imaging. Only sulfonation of thioguanine (lane 22) was not detected.

**TABLE 1**
Sulfonated FDA-approved drugs

| Drug                  | SULT1A1 | SULT2A1 |
|-----------------------|---------|---------|
| Aminoglutethimide     |         |         |
| Arbutamine            |         |         |
| Benindione            |         |         |
| Benzaldehyde          |         |         |
| Bromidin              |         |         |
| Clonidine             |         |         |
| Desvenlafaxine        |         |         |
| Diastereosol          |         |         |
| Edrophonium           |         |         |
| Enoxacin              |         |         |
| Fenoterol             |         |         |
| Furosemide            |         |         |
| Levallophan           |         |         |
| Levonorethodrine      |         |         |
| Mimosine              |         |         |
| Nalbuphine            |         |         |
| Naloxone              |         |         |
| Pentostatin           |         |         |
| Pyrimethamine         |         |         |
| Tizanidine            |         |         |
| Thioguanine           |         |         |

**FIGURE 3. Reactivity of thioguanine.** The figure shows the 35SO4 and 35PAPS bands on a TLC sheet spotted with reactions that either did (+) or did not (−) contain thioguanine. Reaction conditions were as follows: SULT1A1 (25 μM), thioguanine (0 or 1.2 mM), 35PAPS (3.0 μM, SA 0.69 Ci/mmol), MgCl2 (5.0 mM), and KPO4 (50 mM), pH 7.5, 25 ± 2 °C. Reactions were run for 2.0 min, quenched, and spotted, and reactants were separated using anion exchange TLC (see “Experimental Procedures”). The addition of thioguanine caused a 30-fold increase in the levels of SO4 (2–60%).
Predicted SULT Inhibitors—The distance cutoff strategy
described above predicted 136 SULT1A1 inhibitors, which target
64 different areas of metabolism, and 35 SULT2A1 inhibitors,
which target 19 different areas of metabolism. The accuracy
of these predictions was assessed in several ways. Known
SULT inhibitors were counted as confirmed positives. Often,
the inhibition properties of a compound have not been studied,
but its metabolites are known SULT substrates. In such cases,
the derivatives frequently have acquired hydroxyls at reactive
positions; for example, the oxidation of a ketone to a hydroxyl
or the oxidative insertion of a hydroxyl. Because these modifi-
cations are often innocuous in terms of SULT binding (1), such
precursors can often bind and inhibit the enzyme. Although
knowing that a metabolite derivative is a substrate lends cred-
ence to the prediction that a particular compound is an inhib-
itor, validation requires experimentation. Accordingly, com-
ounds that are metabolized to substrates, but have not been
directly tested as inhibitors, are considered likely positives.
Currently, 17 of the 136 predicted SULT 1A1 inhibitors are
confirmed (12.5%), and an additional 34 (25%) are likely posi-
tives. The remaining compounds have not been tested. Of the
35 predicted SULT2A1 inhibitors, 19 are confirmed (54%),
and an additional 8 (23%) are likely positives. The remaining 8 have
not been tested. The predicted inhibitors of 1A1 and 2A1 are listed in
supplemental Table S2, along with a literature refer-
ence (where appropriate), their predicted binding free energies,
and the SULT form(s) to which they bind.

A Case Study: TCA Inhibition—Tricyclic amines, first dis-
covered for their antihistaminic properties (51), are now widely
prescribed as antidepressants (52). These compounds inhibit
reuptake of serotonin and/or norepinephrine in neuronal syn-
apses. An early screen to identify SULT inhibitors among com-
monly prescribed drugs revealed that two TCAs (amitryptiline
and imipramine) inhibit steroid sulfation in human liver extracts
(51). In agreement with this finding, the in silico screens predict
that six TCAs inhibit SULT1A1 (amitryptiline, amoxapine,
imipramine, nortryptilane, protriptyline, and trimipramine) and
that three of the six (amitryptiline, amoxapine, and protripty-
line) also inhibit 2A1. The algorithms predict that the com-
pounds are competitive inhibitors and will bind more tightly to
1A1. To test this prediction and determine the isozyme speci-
ficity and mechanism of the inhibition, two members of the
intersecting set, protriptyline and amoxapine, were tested in
classical initial rate inhibition studies with SULTs 1A1 and 2A1.
A representative data set is shown in Fig. 5. The competitive
inhibition model provided excellent fits to the data in all cases.

Conclusions—Incorporating recent insights into the molec-
ular basis of SULT binding and reactivity into molecular
dynamics models of SULTs 1A1 and 2A1 has produced in silico
versions of these enzymes that are remarkably capable of iden-
tifying substrates. When used to screen a library of 1,455 struc-
turally diverse compounds (the DrugBank Approved Drug
Library), the models identified substrates with
Kd

and

Ki

values and error estimates are listed in Table 2 and, as pre-
dicted, the affinities for 1A1 are greater than those for 2A1.

The Sulfation of FDA-approved Drugs

FIGURE 4. Testing predicted SULT2A1 substrates. Four FDA-approved
drugs were tested as SULT2A1 substrates. The figure shows a STORM image of
a TLC sheet spotted with reactions that used

35

PAPS to label the putative
substrates. The loading scheme: lane 1, control (− acceptor); lane 2, etono-
gestre; lane 3, proflavine; lane 4, dienestrol (two hydroxyls); lane 5, arbu-
tamine (four hydroxyls). Reaction conditions were: SULT1A1 (50 nM dimer),
PnP (2.0

ΜΜ, 1.0 × Ki

J), amoxapine (amox, 0, 30, 60, or 120 µM), MgCl2 (5.0 mM), KPO4 (50
M), pH 7.2, 25 ± 2 °C. Reaction progress was monitored optically at 405 nm
(PnP

εmax

= 15,500 M

−1 cm

−1, pH 7.2). Solid lines through the data represent

the behavior predicted by a weighted fit of the double-reciprocal data using

a competitive inhibition model. K

J = 36 ± 3.5 µΜ.

FIGURE 5. Amoxapine inhibition of SULT1A1 PnP inhibition. Reactions
conditions were: SULT1A1 (50 nM dimer), PAPS (100 µΜ, 315 × Ki

J), PnP (2.0

ΜΜ, 1.0 × Ki

J), amoxapine (amox, 0, 30, 60, or 120 µM), MgCl2 (5.0 mM), KPO4 (50
M), pH 7.2, 25 ± 2 °C. Reaction progress was monitored optically at 405 nm
(PnP

εmax

= 15,500 M

−1 cm

−1, pH 7.2). Solid lines through the data represent

the behavior predicted by a weighted fit of the double-reciprocal data using

a competitive inhibition model. K

J = 36 ± 3.5 µΜ.
that are largely intractable experimentally. It is now feasible to generate hundreds of thousands of synthetically accessible derivatives of a ligand scaffold in silico (53, 54) and, given the findings in this manuscript, to predict their potential to bind and react with SULTs 1A1 and 2A1. The result of such a study would provide an unprecedented view of the selectivity and isoform specificity of these enzymes. Finally, the pharmaceutical and biotech industries now routinely apply in silico analysis in the early stages of their pipelines in an attempt to predict the ADME (absorption, distribution, metabolism, elimination, in the early stages of their pipelines in an attempt to predict the bioavailability) properties of candidate compounds (2, 55–57) and enhance the efficiency with which drugs are created (39, 58).

The SULT models described herein should be quite helpful in enhancing the efficiency with which drugs are created (39, 58). and toxicity) properties of candidate compounds (2, 55–57) and enhance the efficiency with which drugs are created (39, 58).

The SULT models described herein should be quite helpful in enhancing the efficiency with which drugs are created (39, 58).

Quantitative evaluation of the expression and activity of five major sulfotransferases (SULTs) in human tissues. The SULT “pie.” Drug Metab. Dispos. 37, 2255–2261.

12. Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbouy, E., Einer, R., Guo, A. C., and Wishart, D. S. (2011) DrugBank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.

13. Sun, M., and Leyh, T. S. (2010) The human estrogen sulfotransferase. A half-site reactive enzyme. Biochemistry 49, 4779–4785.

14. Comer, K. A., Falany, J. L., and Falany, C. N. (1993) Cloning and expression of human liver dehydroepiandrosterone sulfotransferase. Biochem. J. 289, 233–240.

15. Cook, I., Wang, T., Falany, C. N., and Leyh, T. S. (2012) A nucleotide-gated molecular pore selects sulfotransferase substrates. Biochemistry 51, 5674–5683.

16. Zhang, H., Varlamova, O., Vargas, F. M., Falany, C. N., and Leyh, T. S. (1998) Sulfuryl transfer. The catalytic mechanism of human estrogen sulfotransferase. J. Biol. Chem. 273, 10888–10892.

17. Cook, I., Wang, T., Almo, S. C., Kim, J., Falany, C. N., and Leyh, T. S. (2013) The gate that governs sulfotransferase selectivity. Biochemistry 52, 415–424.

18. Pedersen, L. C., Petrotschenko, E. V., and Negishi, M. (2000) Crystal structure of SULT2A3, human hydroxyoestrogen sulfotransferase. FEBS Lett. 475, 61–64.

19. Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., Pieper, U., and Sali, A. (2006) Comparative protein structure modeling using Modeller. Curr. Protoc. Bioinformatics, Chapter 5, Unit 5.2.9–5.2.30.

20. Pedersen, L. C., Petrotschenko, E., Shvetsov, S., and Negishi, M. (2002) Crystal structure of the human estrogen sulfotransferase-PAPS complex. Evidence for catalytic role of Ser137 in the sulfuryl transfer reaction. J. Biol. Chem. 277, 17928–17932.

21. Wang, J., Ceplak, P., and Kollman, P. A. (2001) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules. J. Comput. Chem. 21, 1049–1074.

22. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhoef, G., Mark, A. E., and Berendsen, H. J. (2005) GROMACS: Fast, flexible, and free. J. Comput. Chem. 26, 1701–1718.

23. Van Gunsteren, W., and Berendsen, H. (1988) A leap-frog algorithm for stochastic dynamics. Mol. Simulation 1, 173–185.

24. Berendsen, H., Postma, J. DiNola, A., and Haak, J. (1984) Molecular dynamics with coupling to an external bath. J. Chem. Phys. 81, 3684–3690.

25. Hess, B., Henk Bekker, H., Herman, J. C., Berendsen, H. J., and Fraaije, J. G. (1998) LINCS. A linear constraint solver for molecular simulations. J. Comput. Chem. 18, 1463–1472.

26. Xavier, D., Karl, G., Bernhard, J., Dieter, S., von G., W., and Man, M. (1999) Peptide folding. When simulation meets experiment. Angew. Chem. Int. Ed. 38, 236–240.

27. (2011) Approved Drug Products with Therapeutic Equivalence Evaluations, 31st Ed., United States Department of Health and Human Services, Food and Drug Administration, Office of Pharmaceuticals, Silver Spring, MD.

28. Weger, B. D., Brenner-Weiss, G., and Dickmeis, T. (2008) Drugbank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.

29. Weger, B. D., Weger, M., Brenner-Weiss, G., and Dickmeis, T. (2008) DrugBank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.

30. Weger, B. D., Weger, M., Brenner-Weiss, G., and Dickmeis, T. (2008) DrugBank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.

31. Weger, B. D., Weger, M., Brenner-Weiss, G., and Dickmeis, T. (2008) DrugBank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.

32. Weger, B. D., Weger, M., Brenner-Weiss, G., and Dickmeis, T. (2008) DrugBank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.

33. Weger, B. D., Weger, M., Brenner-Weiss, G., and Dickmeis, T. (2008) DrugBank 3.0. A comprehensive resource for ‘omics’ research on drugs. Nucleic Acids Res. 39, D1035–D1041.
The Sulfation of FDA-approved Drugs

34501