CoMFA Modeling of Enzyme Kinetics: K_m Values for Sulfation of Diverse Phenolic Substrates by Human Catecholamine Sulfotransferase SULT1A3

Julius Sipilä,*,† Alan M. Hood,§,‡ Michael W. H. Coughtrie,§ and Jyrki Taskinen†

Division of Pharmaceutical Chemistry, Department of Pharmacy, University of Helsinki, P.O. Box 56 (Viikinkaari 5E), 00014 Helsinki, and Department of Molecular and Cellular Pathology, University of Dundee, Ninewells Hospital and Medical School, Dundee DD1 9SY, UK

Received May 7, 2003

Three-dimensional QSAR models were developed for predicting kinetic Michaelis constant (K_m) values for phenolic substrates of human catecholamine sulfating sulfotransferase (SULT1A3). The K_m values were correlated to the steric and electronic molecular fields of the substrates utilizing Comparative Molecular Field Analysis (CoMFA). The evaluated SULT1A3 substrate data set consisted of 95 different substituted phenols, catechols, catecholamines, steroids, and related structures for which the K_m values were available. The data set was divided in three different subgroups in the initial analysis: (1) for the first CoMFA model substrates with only one reacting hydroxyl group were selected ($n = 51$), (2) the second model was build with structurally rigid substrates ($n = 59$), and (3) finally all substrates of the data set were included in the analysis ($n = 95$). Substrate molecules were aligned using the aromatic ring and the reacting hydroxyl group as a template. After the initial analysis different substrate alignment rules based on the existing knowledge of the SULT1A3 active site structure were evaluated. After this optimization a final CoMFA model was built including all 95 substrates of the data set. Cross-validated q^2 values (leave-one-out and leave-n-out) and coefficient contour maps were calculated for all derived CoMFA models. All four CoMFA models were statistically significant with q^2 values up to 0.624. These predictive QSAR models will provide us information about the factors that affect substrate binding at the active site of human catecholamine sulfotransferase SULT1A3.

INTRODUCTION

Sulfation is a major metabolic pathway in detoxification and bioactivation of xenobiotics and endogenous compounds from bacteria to human.[1,2] Sulfuryl transfer from the ubiquitous donor 3′-phosphoadenosine 5′-phosphate (PAPS) to the nucleophilic acceptor group of the substrate is catalyzed by cytosolic and membrane-bound sulfotransferase (SULT) enzymes. There are at least 60 mammalian and avian genes which are known to encode cytosolic sulfotransferases, and many of these proteins have been characterized.[1] All these genes belong to a single gene superfamily which is divided into six families based on similarities in their amino acid sequences and catalytic properties. X-ray crystallographic structures of five mammalian SULTs have been published, including mouse estrogen sulfotransferase mEST,[3,4] human dopamine/catecholamine sulfotransferase SULT1A3,[5,6] human hydroxysteroid/DHEA sulfotransferase SULT2A3,[7,8] human estrogen sulfotransferase SULT1E1,[9] and human phenol sulfotransferase SULT1A1.[10] The structure and function of these enzymes that share several common structural features have recently been reviewed.[1,11]

One of the most important human sulfotransferase enzymes is the catecholamine sulfotransferase (SULT1A3) also known as monoamine sulfotransferase (M-PST), thermolabile phenol sulfotransferase, or human aryl sulfotransferase 3 (HAST3). This enzyme has not been described in any other species to date. Despite its broad substrate spectrum SULT1A3 shows exceptionally high specificity toward catecholamines which is reflected in the high amounts of circulating catecholamine sulfates in humans compared with other species.[1,12] The published crystal structures[5,6] of SULT1A3 are incomplete with numerous missing amino acid residues and no substrate bound at the active site. Site-directed mutagenesis and kinetic studies have revealed several important amino acids which are major determinants of SULT1A3's catalytic activity and substrate specificity.[5,13−16] The sulfuryl transfer mechanism and transition state structure which are common to all cytosolic SULTs have also been studied using X-ray crystallography.[3] Based on these studies it has been concluded that the substrate recognition of SULT1A3 mainly depends on the complementary shape of the substrate and the enzyme's active site, although specific ionic interactions are also important. Traditional two-dimensional QSAR methods have previously been implemented to study structural factors affecting SULT1A3-catalyzed sulfation.[5,17] There is one recently published three-dimensional QSAR study of a related rat aryl sulfotransferase,[18] but so far there are no published three-dimensional QSARs for human SULTs.

The aim of this paper is to present a 3D QSAR model of SULT1A3 utilizing comparative molecular field analysis (CoMFA). The data set used to build the model consists of

† University of Helsinki.
‡ Present address: The Burdock Group, Vero Beach, FL 32962.
§ University of Dundee.
* Corresponding author phone: +358 9 191 59191; fax: +358 9 191 59556; e-mail: julius.sipila@helsinki.fi.

10.1021/ci034089e CCC: $25.00
Published on Web 07/15/2003

Table 1. Statistics of CoMFA Models A−D[a]

model	N	q^2	S_{PRESS}	C	F	SE	r^2
A	51	0.55	0.57	3	78	0.35	0.83
B	59	0.62	0.53	4	87	0.32	0.87
C	95	0.55	0.56	5	103	0.32	0.85
D	95	0.61	0.53	6	101	0.30	0.87

[a] N = number of training set compounds, q^2 = cross-validated (leave-one-out) predictive r^2 of the model, S_{PRESS} = standard error of prediction, C = optimum number of PLS components, $F = F$ value of the model, SE = standard error of the model, r^2 = conventional r^2 of the model.

Table 2. Leave-n-Out Cross-Validation of Models A, B, and C[a]

model	N	mean	SD
A	5	0.50	0.085
	10	0.53	0.033
B	5	0.61	0.054
	10	0.64	0.038
C	5	0.54	0.043
	10	0.55	0.030
D	5	0.60	0.043
	10	0.62	0.023

[a] N = number of cross-validation groups, mean = mean cross-validated q^2, SD = standard deviation of q^2 values. All runs were repeated 100 times.

a large spectrum of SULT1A3 substrates for which the affinity was expressed in terms of enzyme kinetic Michaelis constant (K_m). Existing knowledge of SULT1A3 active site structure was used in development of the model to enhance the relative alignment rules of the substrates. Comparison of the model with previously solved X-ray structures and QSAR models provides us detailed information on factors that affect binding of substrates at the active site of SULT1A3.

MATERIAL AND METHODS

Substrates and Reagents. The evaluated SULT1A3 substrate data set and corresponding apparent Michaelis constant (K_m) values are listed in Table 3.

Enzyme Kinetic Parameters. The K_m values for a part of the substrate set used in this study have been determined earlier[5] in our laboratory with the same methods as described here for the other compounds. The enzyme kinetic parameters were determined using recombinant human SULT1A3 expressed in *Escherichia coli* and purified as described previously.[16] Sulfotransferase activity was determined using a modified PAP^{35}S method described by Foldes and Meek.[19] Purified SULT1A3 was used for estimation of kinetic parameters (K_m and V_{max}) toward various catechols and phenols. All enzyme assays were performed in a final volume of 160 μL containing 6 mM potassium phosphate (pH 6.8), 20 μL of substrate (0.1−3000 μM), 20 μL of PAPS (10 μM final concentration) containing 0.04 μCi PAP35S, and 20 μL of enzyme protein (0.2 μg SULT1A3). Blank reactions contained 20 μL of water in place of substrate. All enzyme reactions were incubated at 35 °C and then terminated by the addition of 200 μL of barium acetate (0.1 M), 200 μL of barium hydroxide (0.1 M), and 200 μL of zinc sulfate (0.1 M). The reaction mixtures were centrifuged at 14 000*g* for 4 min, and 500 μL of supernatant was mixed with 4 mL of scintillation fluid. Radioactivity was quantified by liquid scintillation spectrometry. Kinetic parameters (K_m and V_{max}) were determined by hyperbolic regression analysis with the Hyper.exe software package (Dr. J. S. Easterby, University of Liverpool, UK).

Computational Methods. All computations were performed on SGI Octane workstations using molecular modeling software packages SPARTAN (version 5.0, Wavefunction Inc., Irvine, CA) and SYBYL (version 6.8, Tripos Inc., St. Louis, MO) for building minimum energy conformations of substrate molecules and CoMFA calculations, respectively. Substrate molecular conformations were optimized to the semiempirical AM1 level. Atomic partial charges were calculated according to the Gasteiger−Hückel method. To automate the calculations several Sybyl Programming Language (SPL) scripts were developed.

CoMFA Alignment Rules. The broad structural diversity among the compounds in the original data set and lack of information about the actual binding conformations makes the alignment procedure a critical step in the current study. In the initial analysis the substrates were divided into three different data sets for which three models (A−C) were developed: (A) substrates with only one equivalent phenolic hydroxyl group ($n = 51$); (B) substrates with rigid structure ($n = 59$); and (C) all substrates in the data set ($n = 95$).

The initial CoMFA alignment was based on RMSD fit of one randomly selected phenolic hydroxyl group and the aromatic carbons in positions 1, 3, and 5 of the substrate phenol ring. After the initial analyses the substrate conformations and alignments were refined based on the X-ray crystallographic data of SULT1A3 structure. Some assumptions of active conformations and mutual alignment of substrates were made to achieve the final model: (1) For ethylamines, side chain nitrogen atom position should make the important electronic interaction between the amine and Glu146[5,16] possible. (2) There should be no obvious overlaps of substrate and enzyme side chains when the aligned agglomerate is transferred into the active site of SULT1A3 crystal structure. (3) Relatively small hydrophobic substituents in *ortho* and *meta* positions in respect of the reacting hydroxyl group will be aligned toward the hydrophobic pocket formed by the residues Tyr169, Tyr139, Pro47, and Phe142 in the SULT1A3 structure (Figure 1). (4) The hydrogen atoms of the reacting hydroxyl groups should point toward the catalytic base His108. (5) In case of catechol structure, the hydroxyl group not acting as sulfuryl acceptor should be aligned away from His108 as suggested by Dajani et al.[5] To fulfill these alignment rules the aligned substrates from model C were transferred to the SULT1A3 crystal structure and refined by adjusting the alignment or conformation of substrates when needed. In case of ethylamines the structures were reoptimized holding nitrogen atoms fixed at a position where a positive charge interaction between the amine group and Glu146 is possible. The $\alpha-\beta$ bond of the ethylamine side chain lies perpendicular to the benzene ring in this conformation, as suggested in a previous modeling study.[5] After this refinement model D was obtained with all 95 substrates of the data set. The initial and refined alignments of all molecules are visualized in Figure 2.

CoMFA Parameters. Tripos standard steric and electrostatic fields were used in all CoMFA calculations. H-bonding fields as implemented in SYBYL 6.8 were also evaluated. Standard parameters were used for all fields: steric and

Table 3. Comparison between Actual and Fitted Values of the Michaelis Constant for a Set of Phenolic Substrates of SULT1A3[a]

compound	actual log(1/K_m)	predicted log(1/K_m)	compound	actual log(1/K_m)	predicted log(1/K_m)
1-bromo-3,4-dihydroxy-5-methoxybenzene	5.32	5.06	4-hydroxyestrone	4.48	4.42
2,3-dihydroxybenzaldehyde	5.18[b]	4.72	4-iodophenol	3.78[b]	3.58
2,3-dihydroxybenzoic acid	4.50	4.53	4-isopropylcatechol	4.47[b]	4.17
2,3-dihydroxynaphthalene	5.35[b]	4.79	4-isopropylphenol	2.68[b]	3.37
2-benzylphenol	3.94	4.03	4-methoxyphenol	2.97[b]	3.29
2-bromophenol	5.59	4.92	4-methylcatechol	4.66[b]	4.38
2-chlorophenol	5.33	4.90	4-methylphenol	3.32[b]	3.68
2-fluorophenol	5.16	4.66	4-N-butylphenol	3.55[b]	3.30
2-hydroxyacetophenone	3.05	3.14	4-N-propylphenol	2.86[b]	2.99
2-hydroxybenzaldehyde	4.40	4.00	4-nitrocatechol	4.93[b]	4.84
2-hydroxybiphenyl	4.58	4.67	4-nitrophenol	2.58[b]	3.10
2-hydroxyestradiol	5.22[b]	5.47	4-phenylazophenol	3.91[b]	3.84
2-hydroxyphenylacetic acid	5.52	5.50	4-phenylphenol	3.52[b]	3.44
2-isopropylphenol	4.77	4.97	4-sec-butylphenol	3.67[b]	3.44
2-methoxyphenol	5.19	5.01	4-*tert*-butyl-5-methoxycatechol	4.27	4.11
2-methylphenol	4.86	4.95	4-*tert*-butylcatechol	4.95[b]	5.05
2-nitrophenol	5.17	4.75	5-hydroxydopamine	5.07	5.68
3,4-dihydroxyacetophenone	4.78[b]	4.94	6-hydroxydopamine	5.07	5.26
3,4-dihydroxybenzoic acid	4.80[b]	4.89	apomorphine	5.05[b]	4.87
3,4-dihydroxyphenylacetic acid	4.24	4.19	butyl-4-hydroxybenzoate	4.70[b]	4.55
3,4-dihydroxyphenylalanine	4.82[b]	4.69	carbidopa	4.39	4.36
3,4-dihydroxyphenylglycol	4.48[b]	4.58	catechin	5.13	5.13
3-bromophenol	4.82	4.39	catechol	4.71[b]	4.98
3-chlorocatechol	5.46[b]	5.77	dihydrexidine	6.24	6.33
3-chlorophenol	4.44	4.30	dobutamine	5.33[b]	5.36
3-fluorocatechol	5.19[b]	5.38	dopamine	5.66[b]	5.26
3-fluorophenol	4.12	3.99	epinephrine	5.24	5.61
3-hydroxyacetophenone	3.79	3.45	ethyl-3,4-dihydroxybenzylidenecyanoacetate	5.12	5.35
3-hydroxybenzaldehyde	3.87	3.99	ethyl-4-hydroxybenzoate	4.00[b]	3.92
3-hydroxybiphenyl	4.56	4.39	hydroquinone	3.04	3.43
3-hydroxyphenylacetic acid	3.93	4.13	isoprenaline	5.51	5.45
3-isopropylphenol	4.96	4.46	metanephrine	5.19	5.32
3-methoxycatechol	4.96[b]	5.71	methyl-4-hydroxybenzoate	3.53[b]	3.68
3-methoxyphenol	4.06	4.22	methyldihydroxytetrahydroisoquinoline	5.35[b]	4.76
3-methoxytyramine	5.36	5.14	methylsalicylate	3.33	3.61
3-methylphenol	3.82	4.37	norepinephrine	5.64[b]	5.67
3-nitrophenol	3.94	4.10	normetanephrine	5.12	5.00
4-bromophenol	3.83[b]	3.59	phenol	3.56[b]	4.24
4-bromoresorcinol	5.03	4.99	phenylsalicylate	3.23	3.35
4-chlorocatechol	5.04[b]	5.09	propyl-4-hydroxybenzoate	4.52[b]	4.28
4-chlorophenol	3.89[b]	3.58	pyrogallol	5.62	5.48
4-chlororesorcinol	4.96	5.15	resorcinol	3.96	4.19
4-cyclopentylphenol	3.70[b]	3.87	SKF 3839	4.93	4.90
4-ethylphenol	3.06[b]	3.17	tetrachlorocatechol	5.11[b]	5.28
4-ethylresorcinol	3.66	3.53	tolcapone	4.53	4.77
4-fluorophenol	3.60[b]	3.70	tyramine	4.68[b]	4.44
4-hydroxybenzylcyanide	2.39[b]	2.60	tyrphostin	4.66	4.91
4-hydroxyestradiol	4.87	4.47			

[a] Predicted log(1/K_m) values are from the final model D. [b] Values marked by an asterisk have been published by Dajani et al.[5]

electrostatic potentials were calculated with a +1 charged sp^3 carbon probe, smooth transition, and a cutoff value of 30 kcal/mol. The CoMFA grid spacing was at 2 Å in all calculations, as lowering the grid spacing from this frequently used value did not improve the correlations in the initial studies. Michaelis constant values conversed to log(1/K_m) were used as dependent variable column. To optimize the orientation of the substrates to the CoMFA grid an all space search technique described elsewhere in detail[20] was used. In all space search procedure the aligned substrate molecules are systematically rotated and translated in three-dimensional space to obtain the most consistent model. In this analysis the rotational space was evaluated rotating the substrates around x-, y-, and z-axes by an increment of 30 degrees. Translational space was analyzed translating the substrates 2 Å in each direction by an increment of 0.2 Å. In these orientation/translation analyses the column filter was set at 1 kcal/mol, and the models with highest cross-validated q^2 (leave-one-out) values were selected for further validation.

PLS (Partial Least Squares) Model Validation. After all space search optimization the selected models were validated using leave-one-out and leave-n-out cross-validation procedures. In leave-one-out cross-validation the PLS model is rebuilt omitting each compound of the data set one at a time. This model is used to predict the activity of the omitted compound, and when the procedure is repeated through the whole data set cross-validated r^2 (q^2) values and standard error of predictions can be calculated. In leave-n-out cross-validation n compounds are omitted from the analysis each time when rebuilding the PLS models. Leave-n-out cross-validation was carried out for all PLS models using 5 and 10 randomly selected cross-validation groups. The procedure was repeated 100 times and mean q^2 values

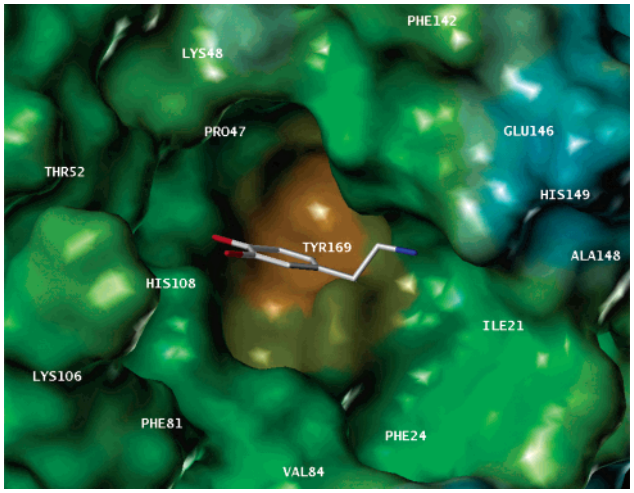

Figure 1. Lipophilic potential mapped to the solvent accessible Connolly surface in the active site of SULT1A3 crystallographic structure.[5] Dopamine is modeled into the active site according to the position of 17β-estradiol in the X-ray crystallographic structure of mouse SULT1E1.[4] The hydrophobic pocket formed by residues Tyr169, Tyr139, Pro47, and Phe142 is clearly visible.

were reported. All validation runs were computed using no column filtering and with CoMFA standard preanalysis scaling.

Final PLS Model. After validation runs a final CoMFA analysis with no validation was performed to obtain graphical contour coefficient maps and conventional r^2 and standard error values.

RESULTS

CoMFA models were derived for two subgroups of substrates, monophenolic compounds (model A) and rigid compounds (model B), and for the whole substrate set (model C). A second model (D) for the whole data set was developed using refined alignment utilizing the published X-ray crystal structure of the enzyme. All derived CoMFA models showed satisfactory statistical characteristics with cross-validated q^2 values (leave-one-out) greater than 0.55 (Table 1). All CoMFA models showed robust behavior in leave-n-out cross-validation analysis with 5 and 10 randomly selected cross-validation groups (Table 2). Using 5 cross-validation groups 20% of data set compounds are randomly omitted from the analysis in each run, which is an extensive test of the predictive capabilities of the QSAR model. The highest correlation between predicted and actual values was achieved with model B ($q^2 = 0.624$), where only relatively rigid compounds were included.

The predicted $\log(1/K_m)$ values vs experimental values of models A−D are plotted in Figure 3. The refined model D showed good fit for all 95 compounds with no outliers ($r^2 = 0.874$, SE = 0.301). In cross-validation runs for model D the difference between actual and predicted K_m values was higher than one logarithmic unit for three compounds: dihydrexidine, 4-hydroxybenzylcyanide, and tolcapone. The fitted non-cross-validated values were within one logarithmic unit in all cases except tyramine in model A.

Graphical View of Developed CoMFA Models. After validation runs a final PLS-analysis was made with no cross-validation to get a graphical view of substrate molecular areas affecting binding to the enzyme (Figure 4).

DISCUSSION

CoMFA is a powerful QSAR tool that has been successfully implemented in various applications for characterization of structural features that affect ligand binding to a given protein.[21] The most critical and demanding phase of any CoMFA is the building of active conformations following with their mutual alignment. In the case of SULT1A3 and the heterogeneous substrate set used in this study there are many possible approaches to the alignment problem. Many of the molecules in the data set possess more than one potential sulfuryl acceptor which also means many possible binding modes at the active site of the enzyme.

In the initial CoMFA we used a simple superimposition procedure without any consideration of SULT1A3 structure or sulfuryl transfer mechanism. Studying the results of the three initial models built with different substrate subsets revealed that special care must be taken for conformational factors when dealing with substrates with flexible side chains. Model A which included all substrates with only one sulfuryl acceptor group showed significantly weaker predictive power than model B built with relatively rigid substrates. In the case of SULT1A3 CoMFA model development finding

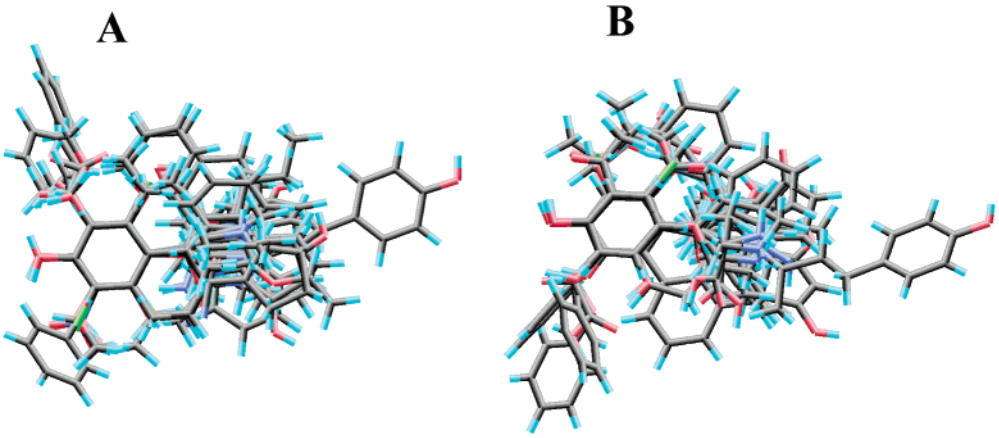

Figure 2. Alignment of all 95 substrate molecules in the data set. (A) In this initial alignment (models A−C) the reacting hydroxyl oxygen and benzene ring carbons are superimposed. (B) Final alignment (model D) that was achieved superimposing reacting hydroxyl groups and benzene ring carbons and directing possible amine groups toward Glu146. Relatively small substituents in the *ortho*-position of the reacting hydroxyl group are also aligned toward the hydrophobic pocket in the SULT1A3 active site.

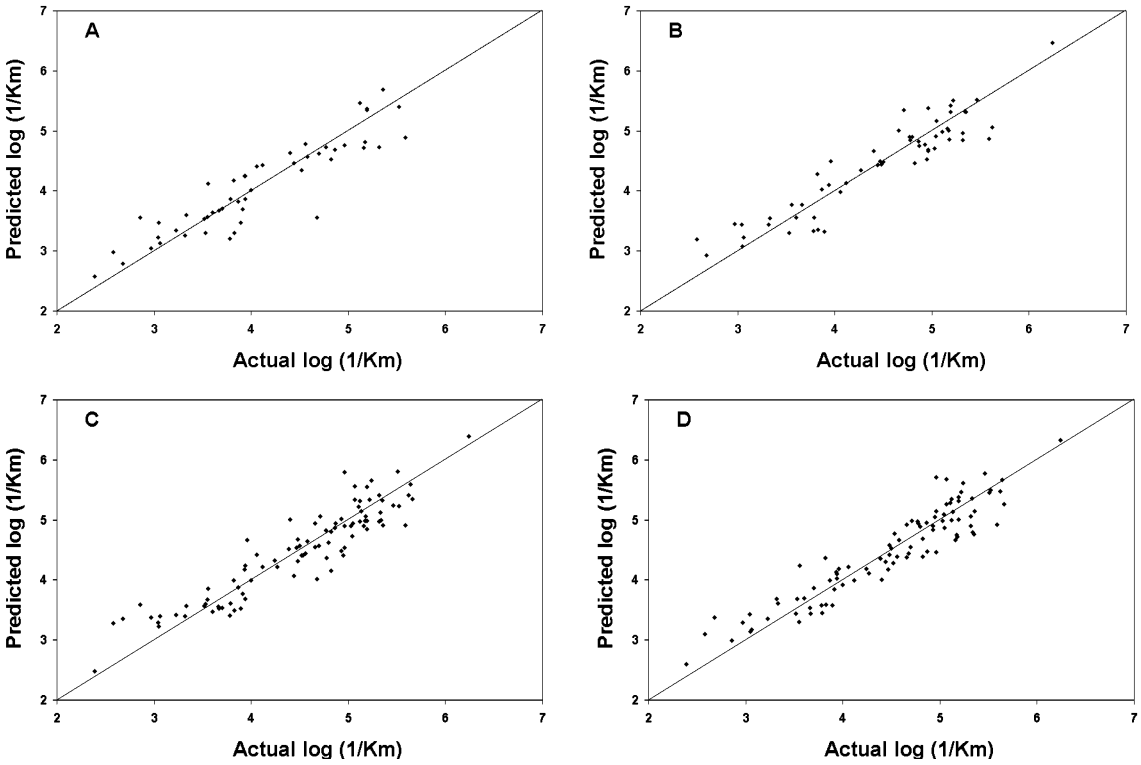

Figure 3. Residual plots of models A−D.

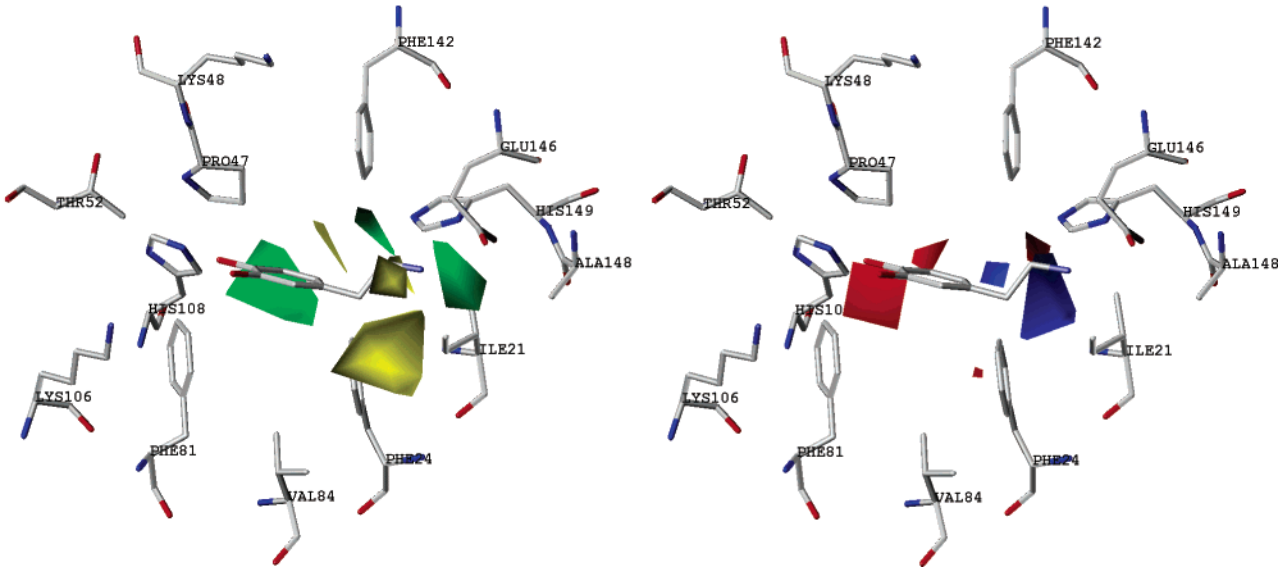

Figure 4. Coefficient contour maps of steric and electrostatic interactions from model D at the active site of SULT1A3 with dopamine as substrate. Green and yellow colors denote favorable and unfavorable regions for bulky substituents, respectively. Positive charge is favored near red regions and negative charge near blue-colored region.

active conformations of substrates with flexible substituents seems to be more crucial than which of the possible sulfuryl acceptor groups are superimposed. This correlates with the in vivo observation that both hydroxyl groups of SULT1A3-metabolized catechols and catecholamines can act as sulfuryl acceptors.

However, the initial CoMFA models were also internally predictive and statistically satisfactory. These models were significantly improved by consideration of SULT1A3 active site structure constraints in the alignment procedure. Homology model and X-ray structure based docking approaches in CoMFA alignment have successfully been implemented for the rat aryl sulfotransferase AST IV[18] and many other proteins.[22−25] With no complete X-ray structure for SULT1A3 available direct docking methods are difficult to implement. There is also recent evidence that the human phenol sulfotransferase SULT1A1 (with 93% sequence similarity to SULT1A3) undergoes significant conformational changes when binding different substrates.[10] It has been proposed that these closely related enzymes show a disorder−order transition upon ligand binding, which explains their activity toward structurally diverse compounds. The fact that both published X-ray structures of SULT1A3 show a considerable region of disorder supports this assumption of a fairly flexible active

site which can adapt its conformation to a given substrate. This kind of substrate binding mechanism makes the use of docking methods to a rigid or semirigid protein structure questionable. The method used in this study takes into account some previously discovered important structural aspects of the enzyme, namely the charged Glu146 and the hydrophobic pocket formed by aromatic residues Tyr169, Tyr139, Pro47, and Phe142. Only X-ray crystallographic structure data of the ordered regions of the enzyme where no significant conformational changes are expected was considered during the refinement of the alignment rules. This approach is free of rigid protein structure and homology model related limitations.

Residual plots for models A, B, and C revealed some compounds which seem to be outliers in this data set. The significant difference between fitted and actual K_m for tyramine in model A is possibly due to the underestimation of the positive effect that ethylamine functionality has for the substrate's affinity to SULT1A3 in this model, which excluded all catecholamines of the data set. Other compounds for which the predictions were relatively inaccurate included 2-halogenated phenols and 4-*n*-propylphenol. It is possible that the long and flexible side chain in 4-*n*-propylphenol adversely interacts with the enzyme in a way that the current models cannot account for. The final model D showed good fit of actual and predicted K_m values, but there were a few compounds with considerable errors of prediction in the cross-validation runs. These outliers for which the error of prediction is more than two times the standard error include dihydrexidine which has the lowest K_m value for the whole data set. Although dihydrexidine is a high-affinity substrate of SULT1A3, it shows a moderate sulfation rate. It is possible that dihydrexidine's binding mode at the enzyme's active site differs from other compounds in the data set. Other outliers in the finished model are 4-hydroxybenzylcyanide and tolcapone. In the development of this model neutral molecules were used, and the negative effect of a cyano group near the charged residue Glu146 in the SULT1A3 structure is neglected which may cause overestimation of 4-hydroxybenzylcyanide affinity.

The graphical view of the finished CoMFA model showed a clear correlation between the derived coefficient contour maps and existing knowledge of SULT1A3 (Figure 4). A region of positive steric interaction and regions of electrostatic interactions were observed in proximity to the hydrophobic pocket formed by Tyr169, Tyr139, Pro47, and Phe142. A region where bulky substituents are not favored is located in *para* position of the reacting hydroxyl group, near Phe24 and Ile21. This observation is consistent with the intuitive interpretation of the activity data: phenols with bulky substituents in the *para* position have high K_m values compared with *ortho*- and *meta*-substituted analogues. The derived CoMFA contour maps clearly emphasize the importance of an electronic interaction near Glu146 in SULT1A3 structure. This is consistent with the fact that this amino acid is crucial for the catecholamine selectivity of the enzyme, which has been revealed in previous studies.[11,16]

CoMFA alignment guided by rational consideration of enzyme active site constraints yielded a statistically robust 3D QSAR model which enables affinity predictions of structurally diverse phenolic compounds. The derived CoMFA maps also correlate well with the current knowledge of SULT1A3 crystal structure, substrate selectivity, and factors affecting catalytic efficiency.

ACKNOWLEDGMENT

This work was supported by in part by a grant from the Commission of the European Communities (BMH4-97-2621, to J.T. and M.W.H.C.) and the Finnish Graduate School in Pharmaceutical Research (Universities of Helsinki and Kuopio, Finland). We thank Prof. Luhua Lai (Institute of Physical Chemistry, Peking University, P.R.China) for providing the scripts for the CoMFA all space search calculations.

REFERENCES AND NOTES

(1) Coughtrie, M. W. H. Sulfation through the looking glass − recent advances in sulfotransferase research for the curious. *Pharmacogenomics J.* **2002**, *2*, 297−308.
(2) Glatt, H. Sulfotransferases in the bioactivation of xenobiotics. *Chem.-Biol. Interact.* **2000**, *129*, 141−170.
(3) Kakuta, Y.; Petrotchenko, E. V.; Pedersen, L. C.; Negishi, M. The sulfuryl transfer mechanism. Crystal structure of a vanadate complex of estrogen sulfotransferase and mutational analysis. *J. Biol. Chem.* **1998**, *273*, 27325−27330.
(4) Kakuta, Y.; Pedersen, L. G.; Carter, C. W.; Negishi, M.; Pedersen, L. C. Crystal structure of estrogen sulfotransferase. *Nature Struct. Biol.* **1997**, *4*, 904−908.
(5) 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.; Coughtrie, M. W. X-ray crystal structure of human dopamine sulfotransferase, SULT1A3. Molecular modeling and quantitative structure−activity relationship analysis demonstrate a molecular basis for sulfotransferase substrate specificity. *J. Biol. Chem.* **1999**, *274*, 37862−37868.
(6) Bidwell, L. M.; McManus, M. E.; Gaedigk, A.; Kakuta, Y.; Negishi, M.; Pedersen, L.; Martin, J. L. Crystal structure of human catecholamine sulfotransferase. *J. Mol. Biol.* **1999**, *293*, 521−530.
(7) Rehse, P. H.; Zhou, M.; Lin, S.-X. Crystal structure of human dehydroepiandrosterone sulphotransferase in complex with substrate. *Biochem. J.* **2002**, *364*, 165−171.
(8) Pedersen, L. C.; Petrotchenko, E. V.; Negishi, M. Crystal structure of SULT2A3, human hydroxysteroid sulfotransferase. *FEBS Lett.* **2000**, *475*, 61−64.
(9) Pedersen, L. C.; Petrotchenko, E.; Shevtsov, S.; Negish, M. Crystal structure of the human estrogen sulfotransferase-PAPS complex. Evidence for catalytic role of Ser137 in the sulfuryl transfer reaction. *J. Biol. Chem.* **2002**, *277*, 17928−17932.
(10) Gamage, N. U.; Duggleby, R. G.; Barnett, A. C.; Tresillian, M.; Latham, C. F.; Liyou, N. E.; McManus, M. E.; Martin, J. L. Structure of a Human Carcinogen-converting Enzyme, SULT1A1. Structural and kinetic implications of substrate inhibition. *J. Biol. Chem.* **2003**, *278*, 7655−7662.
(11) Negishi, M.; Pedersen, L. G.; Petrotchenko, E.; Shevtsov, S.; Gorokhov, A.; Kakuta, Y.; Pedersen, L. C. Structure and function of sulfotransferases. *Archives of Biochemistry & Biophysics* **2001**, *390*, 149−157.
(12) Dousa, M. K.; Tyce, G. M. Free and conjugated plasma catecholamines, DOPA and 3-O-methyldopa in humans and in various animal species. *Proc. Soc. Exp. Biol. Med.* **1988**, *188*, 427−434.
(13) Yoshinari, K.; Petrotchenko, E. V.; Pedersen, L. C.; Negishi, M. Crystal structure-based studies of cytosolic sulfotransferase. *J. Biochem. Mol. Toxicol.* **2001**, *15*, 67−75.
(14) Liu, M. C.; Suiko, M.; Sakakibara, Y. Mutational analysis of the substrate binding/catalytic domains of human M form and P form phenol sulfotransferases. *J. Biol. Chem.* **2000**, *275*, 13460−13464.
(15) Brix, L. A.; Barnett, A. C.; Duggleby, R. G.; Leggett, B.; McManus, M. E. Analysis of the substrate specificity of human sulfotransferases SULT1A1 and SULT1A3: site-directed mutagenesis and kinetic studies. *Biochemistry* **1999**, *38*, 10474−10479.
(16) Dajani, R.; Hood, A. M.; Coughtrie, M. W. A single amino acid, glu146, governs the substrate specificity of a human dopamine sulfotransferase, SULT1A3. *Mol. Pharm.* **1998**, *54*, 942−948.
(17) Campbell, N. R. C.; Vanloon, J. A.; Sundaram, R. S.; Ames, M. M.; Hansch, C.; Weinshilboum, R. Human and Rat-Liver Phenol Sulfotransferase − Structure−Activity-Relationships for Phenolic Substrates. *Mol. Pharm.* **1987**, *32*, 813−819.
(18) Sharma, V.; Duffel, M. W. Comparative molecular field analysis of substrates for an aryl sulfotransferase based on catalytic mechanism and protein homology modeling. *J. Med. Chem.* **2002**, *45*, 5514−5522.

(19) Foldes, A.; Meek, J. L. Rat-Brain Phenolsulfotransferase-Partial Purification and Some Properties. *Biochim. Biophys. Acta* **1973**, *327*, 365−374.
(20) Hou, T. J.; Xu, X. J. Three-dimensional quantitative structure−activity relationship analyses of a series of cinnamamides. *Chemom. Intell. Lab. Syst.* **2001**, *56*, 123−132.
(21) Kubinyi, H. Comparative Molecular Field Analysis (CoMFA). *Encyclopedia of Computational Chemistry*; 2003; DOI: 10.1002/0470845015.cca0470845030.
(22) Bernard, P.; Kireev, D. B.; Chretien, J. R.; Fortier, P. L.; Coppet, L. Automated docking of 82 N-benzylpiperidine derivatives to mouse acetylcholinesterase and comparative molecular field analysis with 'natural' alignment. *J. Comput.-Aided Mol. Des.* **1999**, *13*, 355−371.
(23) Schafferhans, A.; Klebe, G. Docking ligands onto binding site representations derived from proteins built by homology modelling. *J. Mol. Biol.* **2001**, *307*, 407−427.
(24) Huang, X. Q.; Xu, L. S.; Luo, X. M.; Fan, K. N.; Ji, R. Y.; Pei, G.; Chen, K. X.; Jiang, H. L. Elucidating the inhibiting mode of AHPBA derivatives against HIV-1 protease and building predictive 3D-QSAR models. *J. Med. Chem.* **2002**, *45*, 333−343.
(25) Liu, H.; Huang, X. Q.; Shen, J. H.; Luo, X. M.; Li, M. H.; Xiong, B.; Chen, G.; Shen, J. K.; Yang, Y. M.; Jiang, H. L.; Chen, K. X. Inhibitory mode of 1,5-diarylpyrazole derivatives against cyclooxygenase-2 and cyclooxygenase-1: Molecular docking and 3D QSAR analyses. *J. Med. Chem.* **2002**, *45*, 4816−4827.

CI034089E