

Structural determinants of species-selective substrate recognition in human and *Drosophila* serotonin transporters revealed through computational docking studies

Kristian W. Kaufmann,¹ Eric S. Dawson,^{2,3} L. Keith Henry,^{4,7} Julie R. Field,⁴ Randy D. Blakely,^{4,5,6} and Jens Meiler^{1,3,4*}

¹ Department of Chemistry, Vanderbilt University, Nashville, Tennessee

² Department of Biochemistry, Vanderbilt University Medical Center, Nashville, Tennessee

³ Center for Structural Biology, Vanderbilt University Medical Center, Nashville, Tennessee

⁴ Department of Pharmacology, Vanderbilt University Medical Center, Nashville, Tennessee

⁵ Department of Psychiatry, Vanderbilt University Medical Center, Nashville, Tennessee

⁶ Center for Molecular Neuroscience, Vanderbilt University Medical Center, Nashville, Tennessee

⁷ Department of Pharmacology, Physiology and Therapeutics, University of North Dakota, Grand Forks, North Dakota

ABSTRACT

To identify potential determinants of substrate selectivity in serotonin (5-HT) transporters (SERT), models of human and *Drosophila* serotonin transporters (hSERT, dSERT) were built based on the leucine transporter (LeuT_{Aa}) structure reported by Yamashita *et al.* (Nature 2005;437:215–223), PDBID 2A65. Although the overall amino acid identity between SERTs and the LeuT_{Aa} is only 17%, it increases to above 50% in the first shell of the putative 5-HT binding site, allowing de novo computational docking of tryptamine derivatives in atomic detail. Comparison of hSERT and dSERT complexed with substrates pinpoints likely structural determinants for substrate binding. Forgoing the use of experimental transport and binding data of tryptamine derivatives for construction of these models enables us to critically assess and validate their predictive power: A single 5-HT binding mode was identified that retains the amine placement observed in the LeuT_{Aa} structure, matches site-directed mutagenesis and substituted cysteine accessibility method (SCAM) data, complies with support vector machine derived relations activity relations, and predicts computational binding energies for 5-HT analogs with a significant correlation coefficient ($R = 0.72$). This binding mode places 5-HT deep in the binding pocket of the SERT with the 5-position near residue hSERT A169/dSERT D164 in transmembrane helix 3, the indole nitrogen next to residue Y176/Y171, and the ethylamine tail under residues F335/F327 and S336/S328 within 4 Å of residue D98. Our studies identify a number of potential contacts whose contribution to substrate binding and transport was previously unsuspected.

Proteins 2009; 74:630–642.
© 2008 Wiley-Liss, Inc.

Key words: sodium and chloride-dependent neurotransmitter transporters; support vector machine substitution sensitivity map; comparative modeling; ROSETTALIGAND; leucine transporter.

INTRODUCTION

As members of the sodium and chloride-dependent neurotransmitter transporter gene family, serotonin (5-HT) transporters (SERTs) carry out the uptake of 5-HT across plasma membranes in the central nervous system, peripheral nervous system, placenta, platelets, and pulmonary system.^{1,2} SERTs are targets of antidepressants and substances of abuse like cocaine and 3,4-methyldioxy-methamphetamine, commonly known as “Ecstasy.”³ Hydrophathy analyses initially suggested that SERTs are integral membrane proteins with 12 α -helices.^{2,4,5} Site-directed mutagenesis and substituted cysteine accessibility method (SCAM) experiments on putative transmembrane TMs and loops have supported this proposal.^{6–10}

Mutagenesis of key residues has provided insight into the structure and function of SERT. Shortening the ethylamine tail of tryptamine by one methylene group (di-

Additional Supporting Information may be found in the online version of this article.
Abbreviations: 3D, three-dimensional; 5-HT, serotonin; dSERT, *Drosophila* serotonin transporter; hSERT, human serotonin transporter; LeuT_{Aa}, leucine transporter; MTS, methanethiosulfonate reagents; NSS, neurotransmitter:sodium symporter; QSAR, quantitative structure activity relationship; rSERT, rat SERT; SCAM, substituted cysteine accessibility method; SERT, serotonin transporter; SVM, support vector machine; TM, transmembrane spanning domain.

Grant sponsor: NIH; Grant numbers: GM 08320-18 (KWK), GM 080403-01 (JM), DA 07390 (RDB,LKH), DA 022378 (LKH), GM 07628-29 (JRF). Grant sponsor: National Research Alliance on Schizophrenia and Depression Young Investigator Award (LKH).

*Correspondence to: Jens Meiler, Department of Chemistry, Vanderbilt University, VU Station B #351822, 7330 Stevenson Center, Nashville, TN 37235-1822. E-mail: jens.meiler@vanderbilt.edu.

Received 29 September 2007; Revised 5 April 2008; Accepted 23 May 2008
Published online 14 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.22178

methyl-tryptamine to gramine) causes a decrease in substrate uptake in rat SERT (rSERT). The addition of a methylene group, via a D98E mutation, restores uptake of gramine to levels expected for dimethyl-tryptamine, suggesting that residue D98 forms a direct (ion pair) interaction with the substrate.¹¹ Chen *et al.*⁸ implicated I172 and I176 in substrate and inhibitor binding through protection of transporter function from inactivation by methanethiosulfonate reagents (MTS) in cysteine mutants of these residues. Several studies have identified amino acid sequence differences among SERT species that confer alternate specificities for substrates and inhibitors.^{3,12–14} Barker *et al.*¹⁵ used human SERT (hSERT) and *Drosophila* SERT (dSERT) chimeras to implicate Y95 in forming part of the recognition site for citalopram and mazindol, two biogenic amine reuptake inhibitors. Adkins *et al.*¹⁴ used the same approach to show the Y95F hSERT mutant exhibits dSERT-like recognition of *N*-isopropyl tryptamine. Henry *et al.*¹² found that the I172 residue in hSERT displays a marked functional divergence with respect to inhibitor but not substrate potencies when the residue is mutated to its dSERT identity (I172M). Although these advances have identified residues involved in 5-HT and antagonist recognition, interpretation of these data would benefit from a three-dimensional (3D) context provided by high-resolution transporter structures.

Comparative models of SERT have been reported that interpret the structure function implications of site-directed mutagenesis data and substituted cysteine accessibility data using Na^+/H^+ antiporter cryo-EM densities and crystal structure as well as the Lac permease crystal structure.^{16–18} However, the low-sequence homology and low-functional correlation of these templates to SERTs limits the predictive power of these models. The recently reported crystal structure for a bacterial Na^+ -dependent leucine transporter (LeuT_{Aa}), a bona fide member of the neurotransmitter sodium symporter (NSS) protein family, represents a critical break-through for the field.^{19,20} The LeuT_{Aa} structure confirms a predicted topology for NSS members consisting of 12 TM spanning α -helices. Unexpectedly, it features two five-helix bundles arranged in an inverted mirror symmetry. The final two helices, TMs 11 and 12, reside peripheral to the core transporter and may participate in homo-oligomerization.²¹ In the crystal structure of LeuT_{Aa}, the substrate leucine is located in a pocket formed by TMs 1, 3, 6, and 8. Notably, unwound regions in the centers of TMs 1 and 6 serve as contact points for the carboxyl and amine groups of leucine. Beuming *et al.*²² refined the primary sequence alignment of LeuT_{Aa} with a large multisequence alignment of eukaryotic and prokaryotic NSS family members' sequences, resulting in an alignment featuring an improved agreement with available biochemical data that underscores the utility of the LeuT_{Aa} structure.

In this manuscript, ROSETTA comparative modeling^{23,24} and docking²⁵ approaches are invoked for their

power in building accurate models of membrane proteins from distant sequence homologs as recently demonstrated with voltage-gated K^+ channels.²⁶ Our approach involves comprehensive high-resolution docking of 5-HT into SERT comparative models based on the LeuT_{Aa} structure. We refrain from using experimental data during model construction to permit rigorous testing of the predictive power of our models using data derived from site-directed mutagenesis, SCAM, and binding affinity experiments.

METHODS

SERT sequence alignment

The alignment used in the comparative models of dSERT and hSERT on LeuT_{Aa} was synthesized from the alignments of Beuming *et al.*²² The adjusted alignment published between LeuT_{Aa} and rSERT was combined with the alignment of the eukaryotic NSS family.

The SERT sequences were divided into TM and binding site regions based on the LeuT_{Aa} crystal structure. TMs 1, 3, 6, and 8 form the core TMs that surround the leucine binding site. First-shell residues are defined as any residues with a C_α atom within 7 Å of the leucine ligand in the LeuT_{Aa} structure. We define the second-shell binding site residues to be all residues with a C_α atom within 12 Å of the leucine ligand in the LeuT_{Aa}.

SERT comparative model construction

The backbone coordinates of the TM helices from the LeuT_{Aa} crystal structure (PDBID 2A65) were retained in the comparative models of dSERT and hSERT. The loop regions were built in ROSETTA using Metropolis Monte Carlo fragment replacement²⁴ combined with Dunbrack cyclic coordinate descent loop closure.²⁷ In short, ϕ - ψ angles of backbone segments of homologous sequence amino acid fragments from the PDB are introduced for the residues in the loops. After the fragment substitution, slight changes in the ϕ - ψ angles are made to close breaks in the protein chain. Side chains for all residues in the protein were built using ROSETTA's Metropolis Monte Carlo rotamer search algorithm.²⁸ Subsequently, the 10 models generated for both dSERT and hSERT were subjected to eight iterative cycles of side-chain repacking and gradient minimization of ϕ , ψ , and χ angles in ROSETTA.

SERT serotonin docking

A conformational ensemble containing 100 conformations of 5-HT was generated using the mmff94 small molecule force field in Molecular Operating Environment, Chemical Computing Group, Montreal, Quebec, Canada. The ensemble contained representatives from the (\pm) gauche and the anti conformations of the ethylamine tail. Each conformation from the ensemble was placed

into both hSERT and dSERT models for docking calculations using ROSETTALIGAND.²⁵ ROSETTALIGAND placed 5-HT in a random orientation inside a 10 Å cube centered at the same depth as leucine in the LeuT_{Aa} structure. ROSETTALIGAND then simultaneously placed side-chain rotamers around the ligand and optimized the ligand pose in a Metropolis Monte Carlo simulated annealing algorithm. The energy function used during the search contains terms for van der Waals attractive and repulsive forces, statistical energy derived from the probability of observing a side-chain conformation in the PDB, hydrogen bonding, electrostatic interactions between pairs of aminoacids, and solvation assessing the effects of both side-chain side-chain interactions and side-chain ligand interactions.²⁵ Approximately 13,000 docked complexes each for hSERT and dSERT were generated.

When docking small molecules into crystal structures, the ROSETTALIGAND energy function reliably identifies the correct binding mode.²⁵ We validated ROSETTALIGAND's performance on the leucine transporter system by docking leucine into the leucine transporter crystal structure (see supplemental information). Inaccuracies inherent in comparative models preclude identification of the native binding mode based solely on the energy function. However when dealing with comparative models, the energy funnel of the correct binding mode is shallower and local minima can have increased depth (see Fig. 1). Nonetheless, the correct binding mode can occupy a minimum in the energy landscape, which we demonstrate for two comparative model docking studies (see supplemental information). As discussed by recent studies, docking to comparative models remains a difficult task; often experiments can be designed that discriminate between alternate docking poses.^{29,30} Docked complexes occupying a physiologically relevant minimum in the energy landscape might then be identified through testing the predictive power of the models using available biochemical data as a filter.

In the current study, structures with the best protein ligand interaction energies were selected in a first filter. A second filter imposed a 3.6 Å distance between the 5-HT amine tail and one of the D98 side-chain carboxyl oxygens. In a third-round filter, binding modes were chosen that were present in both dSERT and hSERT based on the assumption that the 5-HT binding mode is conserved across the two species.

Serotonin analog docking

After identifying a common binding mode for 5-HT in the both hSERT and dSERT models, 5-HT analogs were placed into the ligand binding site while maintaining the putative binding mode of 5-HT. Each of the analogs then underwent Monte Carlo refinement and gradient energy minimization allowing small adjustments in ligand position and side-chain conformations. For each binding

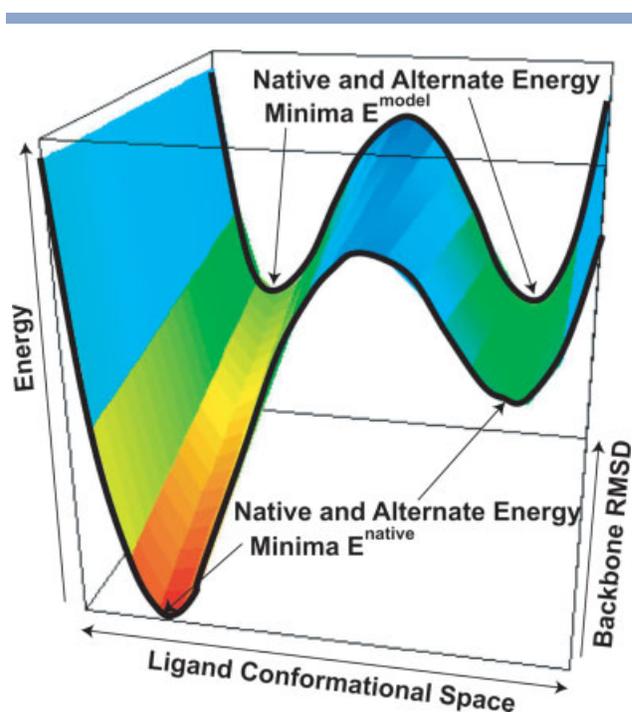


Figure 1

A schematic docking energy landscape is shown as a function of backbone RMSD. The energy is indicated by color from low (red) to high (blue). As the error or RMSD in the backbone increases the native minimum in the energy landscape E^{native} is obscured. Alternate binding modes associated with higher energies can no longer be clearly distinguished from the native binding mode E^{model} . Comparative models by their nature have some error in their atom coordinates. In turn frequently multiple minima are observed when docking small molecules into comparative models. Additional experimental data are required to distinguish between these models.

mode, the nine lowest total ROSETTA energy structures for each analog were selected. Out of the nine structures, the structure with the lowest total ROSETTA energy and with an indole ring less than 1 Å RMSD from the starting position was retained for binding energy calculations. The one exception to the RMSD constraint was the 7-benzyloxy-tryptamine analog, which was allowed to deviate further due to the large bulk of the substitution. The resulting lowest energy structures were visually inspected to verify that they retained the original binding mode.

The binding energy was calculated using,

$$\Delta E_{\text{ligand_binding}} = \Delta E_{\text{protein_bound_state}} - \Delta E_{\text{protein_unbound_state}} \quad (1)$$

where $\Delta E_{\text{protein_unbound_state}}$ is the energy of the protein in the unbound state, and $\Delta E_{\text{protein_bound_state}}$ is the energy of the protein in the bound state plus ligand protein interaction energy. The change in energy, ΔE , is given by

$$\Delta E = \Delta E_{\text{atr}} + \Delta E_{\text{dun}} + \Delta E_{\text{hb}} + \Delta E_{\text{pair}} + \Delta E_{\text{sol}} \quad (2)$$

as was reported previously.^{25,31} ΔE_{atr} is the attractive portion of a van der Waals Lennard-Jones 12-6 potential energy term. ΔE_{dun} is the energy derived from the Dunbrack rotamer probability. ΔE_{hb} is the energy of hydrogen bonds involving side chains. ΔE_{pair} encodes for the energy due to electrostatic interaction between residues. ΔE_{sol} is a Lazaridius-Karplus approximation of the solvation energy. The repulsive portion of the van der Waals energy was removed to decrease noise inherent in the sensitivity of this term. ΔE for each residue was summed to obtain the total ΔE for the protein binding energy. Amino acid residues with a $\Delta E < -1$ were considered to be major contributors to the binding energy.

Model refinement of binding mode with bound Na^+ ion

Molecular models for the sodium (Na^+) ion bound form of both hSERT and dSERT were generated and refined using the following protocol. The ROSETTALIGAND binding mode was taken as the starting point for model refinement using the AMBER force field.³² Briefly, the binding mode models for the hSERT and dSERT were aligned with the published structure of LeuT_{Aa} (PDB ID:2A65) and a single Na^+ ion was added to both models by copying the coordinates of atom Na 752 (Na1 binding site). Models of the hSERT and dSERT sodium ion binding site were then refined with 50 steps of steepest descents and 450 steps of conjugate gradient energy minimization in AMBER9³³ followed by brief (1 ns), low-temperature (50 K) molecular dynamics simulations *in-vacuo* using a distance-dependent dielectric constant, and 12 Å cutoff for nonbonded interactions. Partial charges for 5-HT were developed using the atom-centered point charge method of Bayley *et al.*³³ All other molecular mechanics parameters for 5-HT and ions were taken from the standard AMBER force field. Two-dimensional schematics of the refined hSERT and dSERT ion binding sites were generated with ChemDraw 10.0 (Cambridge Soft) while 3D representations were rendered with PyMol.³⁴

SVM analysis for tryptamine analog pharmacology

Support vector machines (SVM),³⁵ a form of machine learning previously used by this group to study anti-cancer activity of epothilones,³⁶ were applied to derive a substitution sensitivity model for SERT substrates using uptake inhibition data from a previously published study of tryptamine analogs.¹⁴ The freely available software, LIBSVM,³⁷ was applied to 26 tryptamine analogs to derive models for hSERT and dSERT sensitivity to substitution at positions around the indole ring and ethyl amine tail. The binary encoding scheme for each compound was configured to indicate the type of substituent at each of the following positions: R1/2, α , 2, X, R3, 4, 5,

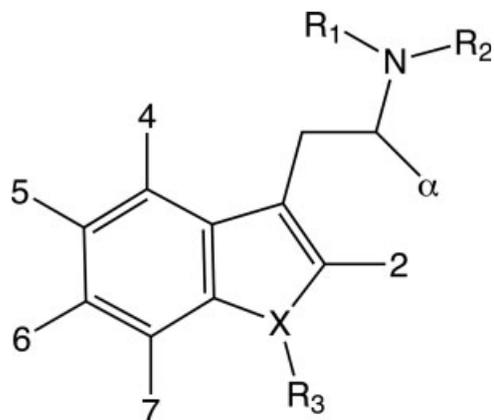


Figure 2

Tryptamine core used in fragment-based substitution encoding for SVM sensitivity maps.

6, 7 (see Fig. 2 and supplemental information). A total of 24 binary inputs are required to uniquely describe the configuration of each of the 26 tryptamine analogs in these nine positions. The resulting input vector of length 24 for each compound is associated with a normalized floating point representation of the experimentally measured binding constant for [³H]5-HT uptake inhibition (K_i) for training of the SVM.

Epsilon support vector regression was applied with a cost of 0.2 and a polynomial kernel function with gamma of 0.1. Optimal cost (c) and gamma (γ) parameters were empirically determined via a systematic search for best RMSD for predicting $\log K_i$ from leave-one-out cross validation. Description of the theory and application of SVM can be found in the following references.^{35,37} The sensitivity to each input was computed as the absolute partial derivative of the output (i.e., SVM-predicted binding constant) with respect to that input. The average sensitivity to substitution was computed by taking the mean of the sensitivities for all inputs coding for substitution at a position on the tryptamine core. The rationale of this approach is that large derivatives identify sensitive inputs that point to more critical regions for binding and vice versa. The average sensitivity to substitution at each position was displayed as a colored molecular surface using PyMOL.³⁴

RESULTS

Our strategy employs comparative modeling, ligand docking, and SAR methodology to address species selectivity for substrate recognition in hSERT and dSERT. Comparative modeling of a target sequence based on a known structural template requires identification of a related structural template, alignment of the target

Table I
Relationship Between Sequence Identity and Expected Model Accuracy

| | | Overall | Loop regions | TMs | Core TMs | 2nd Shell | 1st Shell |
|--|-------|---------|--------------|---------|----------|-----------|-----------|
| Protein sequence identity | hSERT | 17% | 11% | 25% | 35% | 40% | 58% |
| | | 108/630 | 40/362 | 68/268 | 38/108 | 31/77 | 11/19 |
| | dSERT | 18% | 14% | 23% | 33% | 36% | 52% |
| | | 113/622 | 51/354 | 62/268 | 36/108 | 28/77 | 10/19 |
| Expected backbone RMSD to true structure ³⁹ | | >5 Å | >5 Å | ≥2.5 Å | ≈2 Å | – | – |
| Backbone RMSD to LeuT _{Aa} | hSERT | | | 1.6–2.1 | 1.1–1.6 | 1.0–1.3 | 0.9–1.2 |
| | dSERT | | | 1.4–2.3 | 1.1–1.8 | 1.0–1.3 | 0.9–1.2 |

Relationship between sequence identity of hSERT and dSERT to LeuT_{Aa} in specific regions of the protein and the expected model accuracy. Core TMs are TMs 1, 3, 6, and 8. Second shell and 1st shell residues include all residues with C_α atoms within 12 and 7 Å, respectively, of an atom from the leucine ligand in the PDB structure 2A65.

sequence to the structure, model construction, and assessment of the resulting structure.³⁸ Ligand docking programs seek to identify the lowest free energy structure of the ligand–protein complex.³⁹ It is beneficial to categorize the available structural degrees of freedom into ligand internal degrees of freedom (ligand conformation), ligand translation and rotational degrees of freedom (pose), protein side-chain degrees of freedom (rotamer), and protein backbone degrees of freedom. Our approach optimizes all of these degrees of freedom during the course of the model development. In addition, we use SVMs to condense data into substitution sensitivity maps which can be readily compared to the ligand-protein complexes.^{35,36} SVMs allow analysis of data sets containing noise and uneven distribution in the chemical space tested by offering an overview of the available data. The overview can then be interrogated in more depth.

Sequence alignment demonstrates high similarity between the LeuT and the SERT substrate binding sites

Sequence alignments offer insight into the structural similarity of two proteins. The sequence identities in Table I, based on the alignment of hSERT and dSERT to the rSERT–LeuT_{Aa} alignment in Figure 3, reflect regions expected to have different degrees of involvement in the binding of substrates as defined in the Methods section. The sequence identity increases from ~ 15% to greater than 50% as the focus narrows on the first shell of residues in the binding site. As the sequence identity increases, the confidence in the alignment and the resulting quality of the comparative models increases.⁴⁰

SERT comparative models extensively sample backbone and side-chain conformational space

A side by side comparison of hSERT, dSERT, and LeuT_{Aa} models highlight differences that may be responsible for differences in ligand recognition and transport. As can be seen in Figure 4, many side chains of the transporters retain not only their amino acid identity but also the χ angles, supporting the conserved functionality of

these residues. Most of the diversity observed in the binding site is conserved across both dSERT and hSERT and also occurs at the intracellular end of the binding site. The backbone RMSDs in the 20 SERT models range from as little as 0.9 Å in the binding site up to 2.3 Å in trans-membrane spans (see Table I). SCAM accessibility patterns in the regions comprising the binding site show a periodicity that agrees with available experimental data (see Fig. 5).

Serotonin docking comprehensively samples translational and rotational degrees of freedom in the protein–ligand complex and identifies five potential binding modes

Ligand docking searches for the most energetically favorable position of 5-HT in the binding pocket; thus identifying likely structural determinants for 5-HT recognition. Out of the top 100 lowest energy 5-HT complexes for each protein, 22 dSERT models and 24 hSERT models contained a D98 contact. Of those models, six binding modes were present in both proteins. Five of the six binding modes place the amine in approximately the same location as seen for leucine in the LeuT_{Aa} structure. These five modes were carried forward for further analysis and are shown in Figure 6. The first three binding modes Up_a, Up_b, and Up_c have the 5-hydroxyl group oriented in the general direction of the extracellular surface [Fig. 6(a–c)]. In the first binding mode Up_a [Fig. 6(a)], the 5-hydroxyl points toward F335, pushing the phenyl ring of F335 up against the TM 6 helix. The indole nitrogen neighbors T439 in TM 8 at the interface between TMs 3 and 8. For the second binding mode Up_b [Fig. 6(b)], the indole ring is rotated 180° relative to the orientation in Up_a. The indole nitrogen now faces F341. The 5-hydroxyl group is placed against the ring of Y176 lining the upper side of the binding pocket. Up_c [Fig. 6(c)] has the indole ring rotated 90° relative to Up_a. It packs against the phenyl ring of Y176 in a π -stacking interaction. The edge of the ring points toward the interface between TMs 8 and 3, with A173 and G442 opposite to the indole nitrogen in that inter-

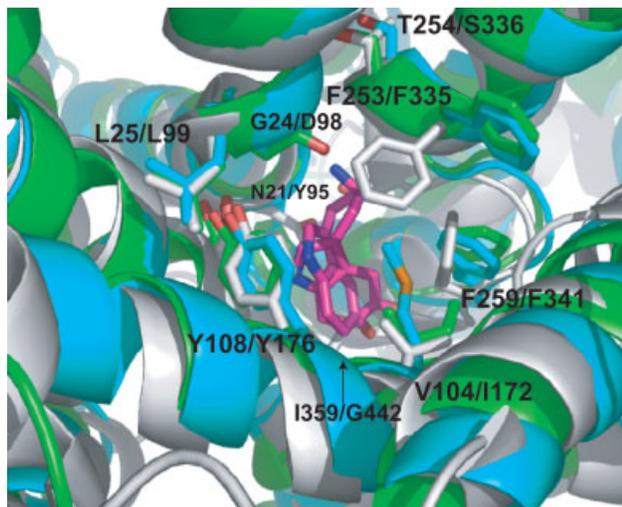


Figure 4

Overlay of hSERT comparative model in green and the dSERT model in cyan on LeuT_{Aa} crystal structure in gray. The conformational space sampled in this study remains close to that of the backbone captured in the LeuT_{Aa} structure. Gradient minimization retains most of the same side-chain interactions, due to the high sequence identity evident in the binding site. This figure was prepared using PyMOL.³⁴

face. In Up_c, the 5-hydroxyl group forms a steric contact with L99. The fourth binding mode (Side) has the 5-hydroxyl bond horizontal in the binding pocket pointing toward T439 and G442 in TM 8 at its interface with TM 3 [Fig. 6(d)]. The indole ring lies sideways in the binding pocket with the side of the indole ring packing against I172. Additionally, the indole nitrogen points toward F335 at top of the binding pocket. The Down binding mode [Fig. 6(e)] shows a 180° rotation of the indole ring relative to the position observed in Up_c. The indole nitrogen is in approximately the same position though pointed more toward T439 and N177. The 5-hydroxyl is now pointed down toward A169 in TM 3 and G342 in TM 6. The residues contributing to the binding energy are boxed in a flattened representation of the binding pocket in each of the five binding modes as shown in Figure 6(I). The agreement of the biochemical data with each of the binding modes is shown in Figure 6(II).

SVM-derived sensitivity maps highlight species differences in the SERT substrate recognition

Adkins *et al.*¹⁴ reported the potencies of 27 tryptamine analogs to inhibit the uptake of [³H]5-HT in the hSERT and the dSERT. Here, we develop SVM-sensitivity maps to visually display differences in the recognition of tryptamine derivatives [Fig. 7(a,b)]. The SVM model trained on tryptamines assayed on the hSERT displays

strong sensitivity to substitution at the 5th position and weaker sensitivity at the R3 indole position and R1 and R2 ethyl amine positions [Fig. 7(a)]. The dSERT SVM model also shows strong sensitivity at the 5th position with a weaker sensitivity at the R3 indole position, the 4 position, and the α-position to the ethyl amine [Fig. 7(b)]. Strong differences in sensitivity between the hSERT and the dSERT SVM maps occur at positions R1, R2, R3, α, 4, 5, and 7 [Fig. 7(c)]. The hSERT SVM map shows higher sensitivity at the R3, 7, R1, R2, and 5 positions in order of increasing difference in sensitivity. The dSERT SVM map shows higher sensitivity at the α, and 4 position in increasing order of difference in sensitivity. Care is taken to avoid over-interpretation of the SVM maps by resorting to the original data when making use of the maps in the context of modeling.

Serotonin analog docking probes ROSETTALIGAND identified binding modes through binding energy prediction

It can be hypothesized that SERTs recognize tryptamine analogs in a conserved manner such that the indole ring occupies the same position in the binding pocket. With this in mind, the native binding mode for 5-HT should explain the differences in the binding affinity seen for other tryptamine analogs. Representative deviations of the indole ring when docking 5-HT analogs in the Down binding mode are shown in Figure 8. In the Down mode, the substitution of the indole nitrogen causes Y176 to change rotamers. Substitutions at the 5th position interact with residues V343, G442, and A169 in

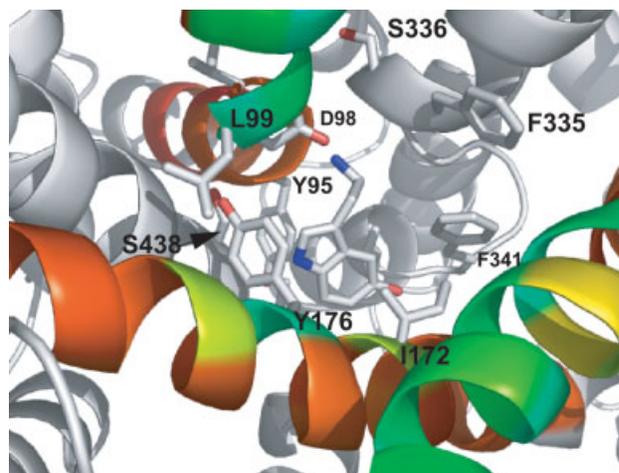


Figure 5

hSERT Down binding mode with substituted cysteine accessibility mapped onto TM 1, 3, and 10. Red to blue scale indicates no sensitivity to large sensitivity to MTS attack of a cysteine substituted at that residue. All three helices show patterns consistent with the helix orientations in the models. This figure was prepared using PyMOL.³⁴

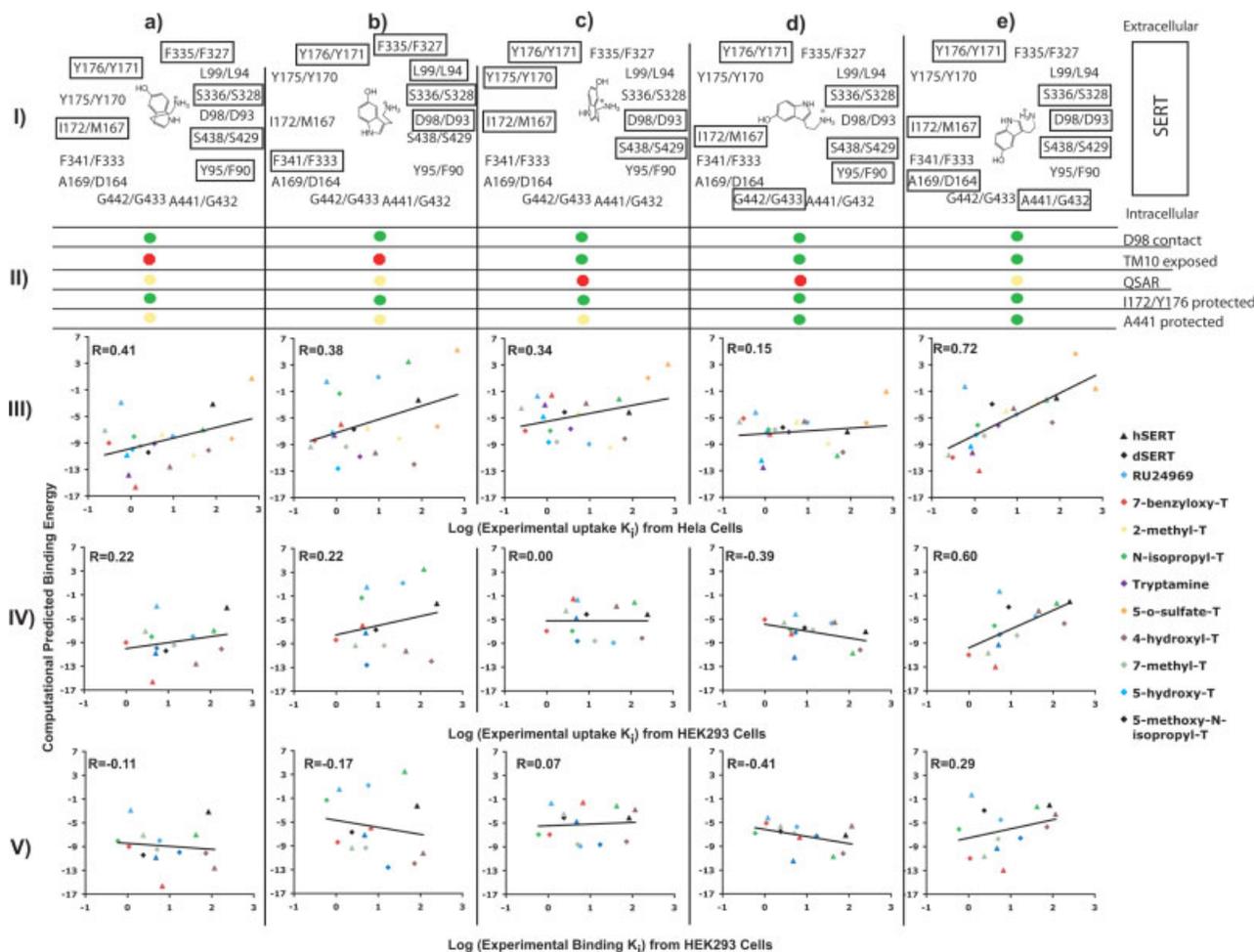


Figure 6

For each of the docked complexes (a) Up_a, (b) Up_b, (c) Up_c, (d) Side, (e) Down (I) shows a flattened representation of the binding site with residues contributing most to the computational binding energy outlined in rectangles with black borders. (II) shows agreement of each docking mode with biological data. Each mode contains a D98 contact. Up_a and Up_b display contacts with TM10 that contradict the lack of protection from MTS inactivation. Up_c and Side binding modes do not match the SVM species difference maps. All the modes show interaction with I172 and Y176 explaining protection against MTS modification. The Side and Down modes pack closely to A441 in a manner which may explain protection of A441C by 5-HT from MTS modification. (III–V) Correlation plots for predicted $\log K_i$ (calculated on computational binding free energy of tryptamine analogs in these modes) and $\log K_i$ for uptake in Hela cells (III), for uptake in HEK293 cells (IV), and for binding in HEK293 cells (V). hSERT values are given in triangles and dSERT values in diamonds. All experimental transport and binding data taken from Adkins *et al.*¹⁴

this binding mode. Figure 6(III–V) shows the correlations of the predicted binding free energies of ligand binding and the log of the uptake and binding K_i values extracted from experimental competitive uptake and binding assays by Adkins *et al.*¹⁴ The Down mode shows the highest correlation for all three datasets. The correlation coefficient of the Down binding mode to the log uptake K_i data from Hela cells is 0.72. The correlation coefficient to log uptake K_i data from HEK293 cells is 0.60. The coefficient falls to 0.29 when compared with log binding K_i data extracted from HEK293 competition binding assays. The first two datasets of uptake K_i in

HEK293 and Hela cells assess the ability of tryptamine analogs to competitively inhibit uptake of tritiated serotonin across membranes with the SERT transporter. The third dataset of binding K_i s assesses the ability of tryptamine analogs to compete with a high-affinity inhibitor to bind to the SERTs. This third category measures a competitive binding event, a more close approximation to the binding energy measured in this study. However, binding is thought to be an important step during transport, and the uptake studies examine the ability of chemical similar compounds to compete. Thus, uptake potency provides a relevant assessment of binding. In any

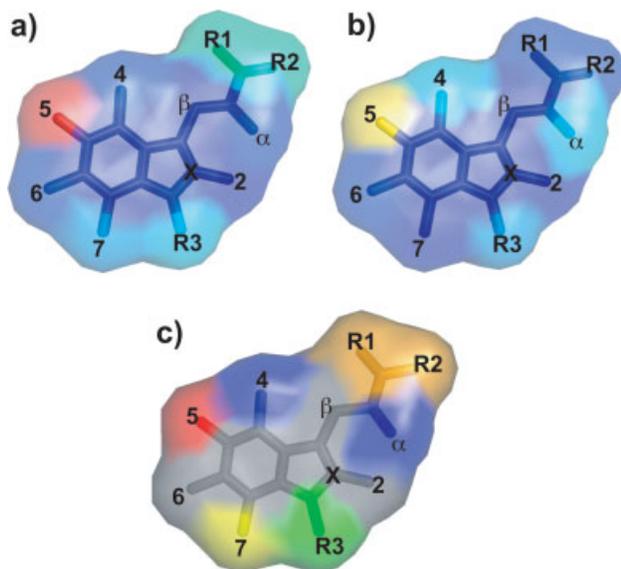


Figure 7

Sensitivities of positions to substitution predicted from support vector machine trained on SERT transporter substrate uptake K_s s. Blue to red gradient indicates low to high sensitivity. (a) hSERT, (b) dSERT, (c) difference map (hSERT-dSERT) of the raw sensitivities. Blue shows higher sensitivity for dSERT. Green to red indicates moderate to higher sensitivity in hSERT. This figure was prepared using PyMOL.³⁴

case, the Down binding mode remains the best correlated of the five binding modes [see Fig. 6(e)].

Model minimization in amber force field confirms hydrogen bonding contacts of 5-OH group

We refined our final models using the AMBER force field employing a short molecular dynamics simulation as a minimization tool.⁴¹ We leverage the ability of the molecular mechanics force field in AMBER to model ligand flexibility to optimize the models for the hSERT and dSERT 5-HT Down binding mode (see Fig. 9). As this calculation is a local refinement with minimal movements, the ROSETTALIGAND conformations are not altered significantly. However, the geometry of hydrogen bonds and other local interactions are improved. The conformation identified by ROSETTALIGAND proves to be stable after 1 ns of molecular dynamics. The overall RMSD of the binding site in both models before and after refinement is <1.0 Å indicating that, even though the sodium ion is not explicitly included in our model building and ligand docking to identify the “Down” binding mode, the conservation of the site implicitly encodes this information. The 5-OH substituent of 5-HT maintains a hydrogen bond to the dSERT D164 side-chain carbonyl oxygen, whereas in the hSERT the 5-OH of 5-HT

forms transient hydrogen bonds to the backbone oxygens of residue A169 (dSERT D164) and A441 (dSERT G432).

DISCUSSION

This study examines two primary questions; “Can docking of 5-HT into comparative models of SERTs identify a physiologically relevant binding mode consistent with known mutagenesis, SCAM, and SAR data?” and “If so, what are the implications for SERT substrate recognition?” Computational docking on its own is unlikely to present a single correct solution due to the errors inherent in comparative models.³⁰ However, docking to comparative models may yield a physiologically relevant binding mode²⁹ (see supporting information). Functional conservation, sequence identity, and biochemical structural data all indicate promising potential for comparative models based on LeuT_{Aa} structure. Chothia and Lesk⁴² found that functional conservation of proteins often implies a higher structural conservation than sequence identity would imply. In a study of comparative modeling for membrane proteins, Forrest *et al.*⁴⁰ reported that sequence identities above 30% in the transmembrane domains yield models with C_{α} -RMSD of ~ 2 Å to the true structure. Biochemical structural information such as the SCAM profiles of TMs 1, 3, and 10 in SERTs are consistent with the LeuT_{Aa} structure.²²

No single model resulting from this process is guaranteed to satisfy all the biochemical data available. How-

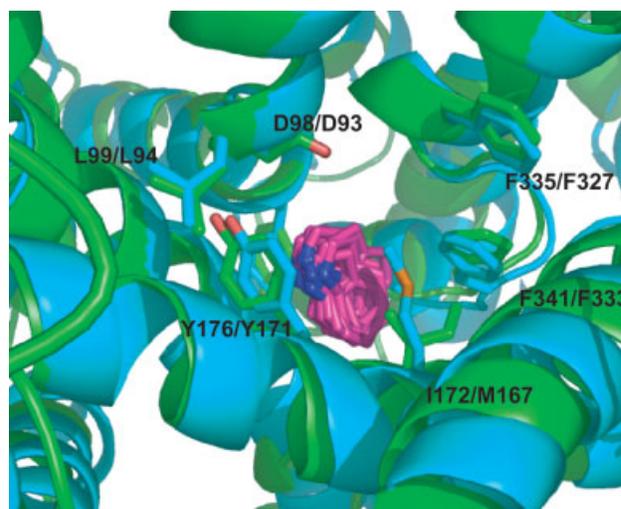


Figure 8

A superimposition of the indole ring of tryptamine derivatives in the Down binding mode is shown for hSERT and dSERT docking. It highlights the conserved manner in which tryptamine derivatives are recognized by SERTs. This figure was prepared using PyMOL.³⁴

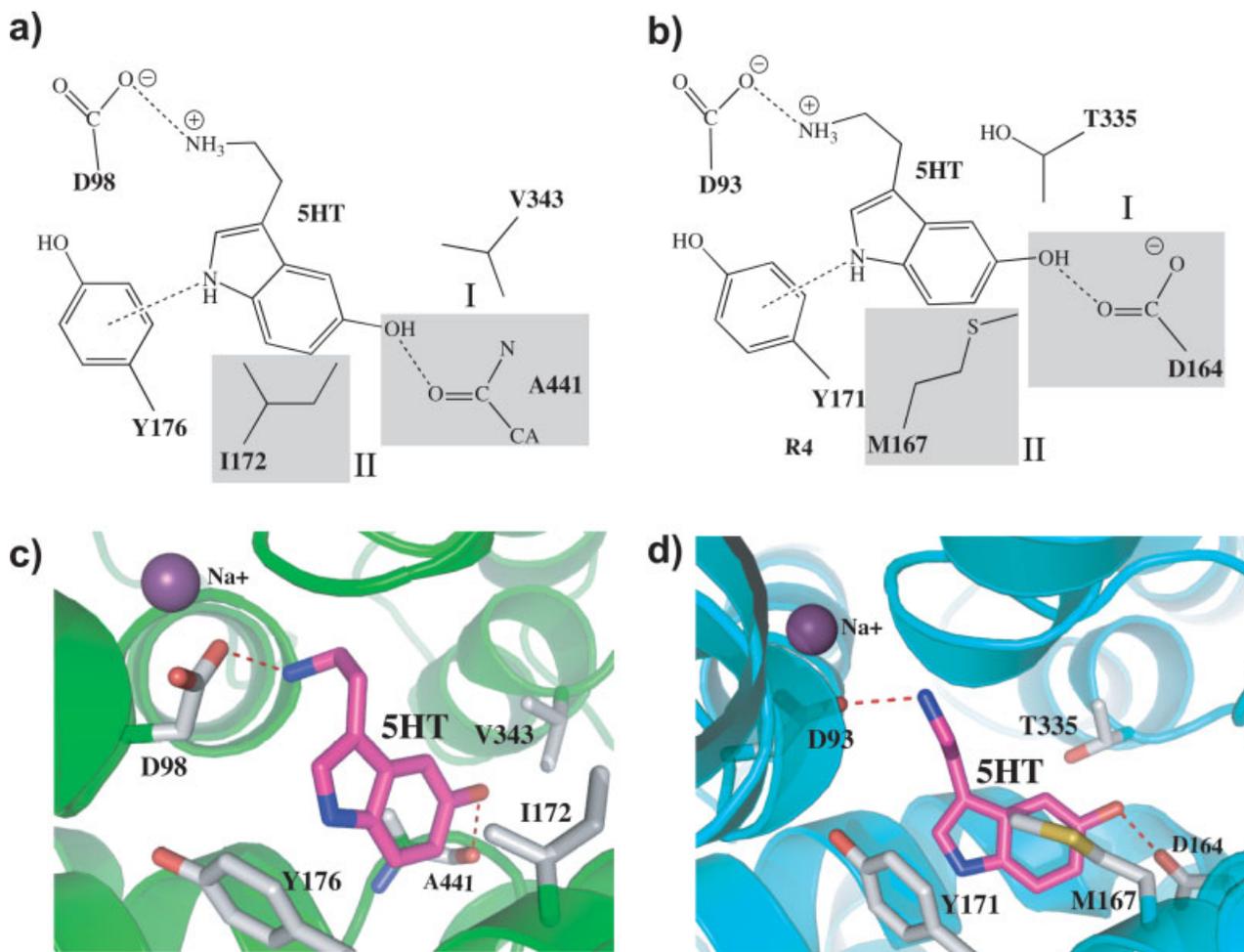


Figure 9

The Down binding mode in the hSERT and dSERT models. Dashed lines in (a) and (b) represent stable hydrogen bonding interactions observed during the 1 ns AMBER refinement of the best ROSETTALIGAND model [Fig. 6(e)] of the substrate binding site. The dashed line from 5-HT to the aromatic ring of Y176 marks a T-type ring stacking interaction. The dashed lines highlight major differences of the hSERT and dSERT models in the substrate binding site: (I) The A441/D164 hydrogen bonding interactions with the 5-OH position of 5-HT. (II) I172/M167 packing interactions with 5-HT indole ring. Panels (c) and (d) show 3D representations of the Down binding mode in hSERT and dSERT models.

ever, in our study unbiased sampling of possible binding modes produced a single binding mode in line with all biochemical data. The collective satisfaction of these constraints indicates the physiological relevance of the Down binding mode shown in Figures 6(e) and 9. For example, in the Down mode residues, I172 and Y176 are protected from MTS modification and subsequent inactivation of transport. Only bulky or charged mutations at I172 have a significant effect on 5-HT transport,¹² indicating a purely steric impact of this position on the binding site as is indicated by the packing against the side of the indole ring. The hSERT G100A mutant is transport deficient but maintains an unperturbed binding affinity.⁴³ Since the Down binding mode lies below G100, G100A would not significantly perturb this binding mode. TM

10 residues cannot be protected from MTS attack and inactivation by 5-HT binding.⁶ The Down binding mode predicts this since it leaves TM 10 amino acids, which are sensitive to MTS modification, solvent accessible. Finally, the A441C mutant is protected from MTS access by 5-HT⁴⁴ inline with the proximity of A441 to the 5-OH group. The sum of all these experimental data points support the Down binding mode as a physiologically relevant placement for 5-HT in the binding site.

SVM sensitivity maps reveal differences in the sensitivities of dSERT and hSERT to substitution at the R3, 4, 5, and α -positions (see Fig. 7). The R3 indole nitrogen displays sensitivity to bulky substituents in hSERT.¹⁴ An isopropyl substitution causes a significant decrease in transport, whereas a methyl substituent in the same posi-

tion causes little difference in uptake. These data indicate that the indole nitrogen likely faces a sterically restricted area in hSERT. The Down binding mode places the indole nitrogen R3 substituents proximal to Y176/Y171. Y176 has been shown to be important for transport⁸; thus, it is not surprising that the substitutions perturbing this residue are detrimental to transport. Adkins *et al.* identified a mutant hSERT, Y95F, which minimizes this effect.¹⁴ Since no direct contact between R3 substituents and Y95 is seen in our models, we hypothesize an indirect effect as follows: the tryptamine *N*-isopropyl substitution causes a shift in the indole ring toward the bottom of the pocket where Y95 is located in hSERT (F90 in dSERT). Mutation at position 95 allows for a structural rearrangement that accommodates additional bulk at the indole nitrogen position. If this is the case, then bulk reducing mutations at neighboring residues, such as V343, L344, and A441, could have a similar effect and could serve to test our hypothesis. In contrast to hSERT, the intracellular base of the binding site in dSERT exhibits a more polarizable nature (e.g., hydrophobic to polarizable I172/M167, V343/T335 hydrophobic to polar, and A169/D164 hydrophobic to charged, see Fig. 9). The hydrogen bond seen between the 5-OH of 5-HT and the side chain of D164 reinforces this view. Furthermore, sensitivity to substitution at positions 4 and 5 as shown in the SVM sensitivity maps agree with the Down binding mode by placing hydroxyl groups near V343/T335 and A169/D164 in the hSERT/dSERT [Figs. 7(c) and 9].

The Down binding mode merits experimental investigation given agreement with the above biochemical data. The difference in polarity in this region in combination with the Down mode placing the 5-OH in this region implies that dSERT and hSERT should exhibit a differential preference for polarity surrounding the 4 and 5 position of the tryptamine ring. Further studies with species switching mutations of the above residues will ascertain the role of these residues in substrate specificity for 4- and 5-position tryptamine derivatives. Since the sparseness in the dataset for substitutions at α , R3, and 4 limits the further analysis of determinants of sensitivities to substitution at these positions, uptake and binding assays experiments with additional substrates modified at these positions should be useful in the context of our models.

The Down binding mode places the indole ring such that the 6 and 7 positions of the tryptamine core point toward the interface between TM 8 and TM 3. The amino acid identities of residues at this interface do not change significantly in hSERT and dSERT. However, future experiments with site-directed mutants in this region may verify the orientation of indole ring of the Down binding mode. One prediction is that a hSERT T439A mutant would display differential recognition of polarity switching substitutions at the 7th position on the tryptamine core. Additional hSERT mutants, such as G442S, A173S, and A169S, would impact recognition of

6-position substituted tryptamines with varied hydrogen bonding capabilities. Assessing the function of these mutants in both hSERT and dSERT backgrounds could validate the assumption of a conserved mode for tryptamines in SERTs. Should the assumption prove incorrect, this constraint on the binding mode selection could be changed to find modes consistent with new experimental findings.

Despite the advances made with the current models, much still remains unknown. The LeuT_{Aa} structure captures but one state in a multistep transport process. Structures of other states in the transport process are needed to fully understand species selectivity for substrates. Additionally, the LeuT_{Aa} structure lacks a chloride in the binding site known to be required for function of the SERT. Studies are forthcoming to elucidate mechanism of chloride coupling in transport.

Jorgensen *et al.*⁴⁵ independently performed a manual docking and molecular dynamics study with 5-HT in hSERT. Interestingly, the binding mode identified is similar to our Down mode. Celik *et al.*⁴⁶ recently reported a study on hSERT using the paired mutant-ligand analog complementation approach. They reported an alternate binding mode using this approach. Our approach places a lower priority on their proposed binding mode as it seems less consistent with the cross-species sensitivities reported in the SVM sensitivity maps. We expect hSERT and dSERT to show differences in the amino acids in regions surrounding the 5th position and the N-position. Of course, hSERT and dSERT could bind 5-HT in different modes, but this is unlikely. Our study applies a different approach of comparing multiple tryptamine derivatives in both hSERT and dSERT, thereby identifying structural determinants of substrate specificity in these transporters.

CONCLUSIONS

Docking of 5-HT into hSERT and dSERT identifies a single conserved binding mode, in which the predicted binding energy of tryptamine derivatives correlates with inhibition uptake constants ($R = 0.72$). The Down binding mode curls the ethylamine tail under F335 and S336 and orients the 5-OH group toward A169 with the indole nitrogen facing the top of the binding site covered by Y176. This binding mode correctly predicts, qualitatively, the decreased modification by SCAM reagents of cysteines substituted at I172, Y176, A441, and the extracellular half of TM 10 due to binding of 5-HT. The mode posits that polarity differences caused by A169D and V343T changes could be responsible for species selectivity observed for hSERT and dSERT recognition of tryptamine derivatives. As additional mutations in SERTs are produced and characterized, particularly in the context of substituted tryptamines, our models should be capable of local refinement to even more precisely focus its utility.

ACKNOWLEDGMENTS

The authors thank David Nannemann and Jarrod Smith for assistance in the development of these models. They also thank the members of the Meiler and Blakely laboratory for helpful discussions.

REFERENCES

- Rothman RB, Baumann MH. Therapeutic and adverse actions of serotonin transporter substrates. *Pharmacol Ther* 2002;95:73–88.
- Ramamoorthy S, Bauman AL, Moore KR, Han H, Yang-Feng T, Chang AS, Ganapathy V, Blakely RD. Antidepressant- and cocaine-sensitive human serotonin transporter: molecular cloning, expression, and chromosomal localization. *Proc Natl Acad Sci USA* 1993;90:2542–2546.
- Roman DL, Saldana SN, Nichols DE, Carroll FI, Barker EL. Distinct molecular recognition of psychostimulants by human and *Drosophila* serotonin transporters. *J Pharmacol Exp Ther* 2004;308:679–687.
- Blakely RD, Berson HE, Fremeau RT, Jr, Caron MG, Peek MM, Prince HK, Bradley CC. Cloning and expression of a functional serotonin transporter from rat brain. *Nature* 1991;354:66–70.
- Hoffman BJ, Mezey E, Brownstein MJ. Cloning of a serotonin transporter affected by antidepressants. *Science* 1991;254:579–580.
- Keller PC, II, Stephan M, Glomska H, Rudnick G. Cysteine-scanning mutagenesis of the fifth external loop of serotonin transporter. *Biochemistry* 2004;43:8510–8516.
- Henry LK, Adkins EM, Han Q, Blakely RD. Serotonin and cocaine-sensitive inactivation of human serotonin transporters by methanethiosulfonates targeted to transmembrane domain I. *J Biol Chem* 2003;278:37052–37063.
- Chen JG, Sachpatzidis A, Rudnick G. The third transmembrane domain of the serotonin transporter contains residues associated with substrate and cocaine binding. *J Biol Chem* 1997;272:28321–28327.
- Chen JG, Liu-Chen S, Rudnick G. Determination of external loop topology in the serotonin transporter by site-directed chemical labeling. *J Biol Chem* 1998;273:12675–12681.
- Zhang YW, Rudnick G. The cytoplasmic substrate permeation pathway of serotonin transporter. *J Biol Chem* 2006;281:36213–36220.
- Barker EL, Moore KR, Rakhshan F, Blakely RD. Transmembrane domain I contributes to the permeation pathway for serotonin and ions in the serotonin transporter. *J Neurosci* 1999;19:4705–4717.
- Henry LK, Field JR, Adkins EM, Parnas ML, Vaughan RA, Zou MF, Newman AH, Blakely RD. Tyr-95 and Ile-172 in transmembrane segments 1 and 3 of human serotonin transporters interact to establish high affinity recognition of antidepressants. *J Biol Chem* 2006;281:2012–2023.
- Rodriguez GJ, Roman DL, White KJ, Nichols DE, Barker EL. Distinct recognition of substrates by the human and *Drosophila* serotonin transporters. *J Pharmacol Exp Ther* 2003;306:338–346.
- Adkins EM, Barker EL, Blakely RD. Interactions of tryptamine derivatives with serotonin transporter species variants implicate transmembrane domain I in substrate recognition. *Mol Pharmacol* 2001;59:514–523.
- Barker EL, Perlman MA, Adkins EM, Houlihan WJ, Pristupa ZB, Niznik HB, Blakely RD. High affinity recognition of serotonin transporter antagonists defined by species-scanning mutagenesis. An aromatic residue in transmembrane domain I dictates species-selective recognition of citalopram and mazindol. *J Biol Chem* 1998;273:19459–19468.
- Ravna AW, Edvardsen O. A putative three-dimensional arrangement of the human serotonin transporter transmembrane helices: a tool to aid experimental studies. *J Mol Graph Model* 2001;20:133–144.
- Ravna AW, Sylte I, Dahl SG. Molecular mechanism of citalopram and cocaine interactions with neurotransmitter transporters. *J Pharmacol Exp Ther* 2003;307:34–41.
- Ravna AW, Jaronczyk M, Sylte I. A homology model of SERT based on the LeuT(Aa) template. *Bioorg Med Chem Lett* 2006;16:5594–5597.
- Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E. Crystal structure of a bacterial homologue of Na⁺/Cl⁻ dependent neurotransmitter transporters. *Nature* 2005;437:215–223.
- Henry LK, Defelice LJ, Blakely RD. Getting the message across: a recent transporter structure shows the way. *Neuron* 2006;49:791–796.
- Just H, Sitte HH, Schmid JA, Freissmuth M, Kudlacek O. Identification of an additional interaction domain in transmembrane domains 11 and 12 that supports oligomer formation in the human serotonin transporter. *J Biol Chem* 2004;279:6650–6657.
- Beuming T, Shi L, Javitch JA, Weinstein H. A comprehensive structure-based alignment of prokaryotic and eukaryotic neurotransmitter/Na⁺ symporters (NSS) aids in the use of the LeuT structure to probe NSS structure and function. *Mol Pharmacol* 2006;70:1630–1642.
- Misura KMS, Baker D. Progress and challenges in high-resolution refinement of protein structure models. *Proteins* 2005;59:15–29.
- Rohl CA, Strauss CE, Chivian D, Baker D. Modeling structurally variable regions in homologous proteins with rosetta. *Proteins* 2004;55:656–677.
- Meiler J, Baker D. ROSETTALIGAND: protein-small molecule docking with full side-chain flexibility. *Proteins* 2006;65:538–548.
- Yarov-Yarovoy V, Baker D, Catterall WA. Voltage sensor conformations in the open and closed states in ROSETTA structural models of K(+) channels. *Proc Natl Acad Sci USA* 2006;103:7292–7297.
- Canutescu AA, Dunbrack RL, Jr. Cyclic coordinate descent: a robotics algorithm for protein loop closure. *Protein Sci* 2003;12:963–972.
- Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. Design of a novel globular protein fold with atomic-level accuracy. *Science* 2003;302:1364–1368.
- DeWeese-Scott C, Moulton J. Molecular modeling of protein function regions. *Proteins* 2004;55:942–961.
- Kairys V, Fernandes MX, Gilson MK. Screening drug-like compounds by docking to homology models: a systematic study. *J Chem Inf Model* 2006;46:365–379.
- Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci USA* 2002;99:14116–14121.
- Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 2000;21:1049–1074.
- Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J Phys Chem* 1993;97:10269–10280.
- DeLano WL. The PyMOL molecular graphics system. San Carlos, CA, USA: DeLano Scientific; 2002.
- Vapnik VN. *Statistical learning theory*. New York: Wiley; 1998. p xxiv, 736 p.
- Bleckmann A, Meiler J. Epothilones: quantitative structure activity relations studied by support vector machines and artificial neural networks. *QSAR Combinatorial Sci* 2003;22:722–728.
- Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- Baker D, Sali A. Protein structure prediction and structural genomics. *Science* 2001;294:93–96.
- Ferrara P, Gohlke H, Price DJ, Klebe G, Brooks CL, III. Assessing scoring functions for protein-ligand interactions. *J Med Chem* 2004;47:3032–3047.

40. Forrest LR, Tang CL, Honig B. On the accuracy of homology modeling and sequence alignment methods applied to membrane proteins. *Biophys J* 2006;91:508–517.
41. Summa CM, Levitt M. Near-native structure refinement using in vacuo energy minimization. *Proc Natl Acad Sci USA* 2007;104:3177–3182.
42. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. *EMBO J* 1986;5:823–826.
43. Kristensen AS, Larsen MB, Johnsen LB, Wiborg O. Mutational scanning of the human serotonin transporter reveals fast translocating serotonin transporter mutants. *Eur J Neurosci* 2004;19:1513–1523.
44. Androutsellis-Theotokis A, Rudnick G. Accessibility and conformational coupling in serotonin transporter predicted internal domains. *J Neurosci* 2002;22:8370–8378.
45. Jorgensen AM, Tagmose L, Jorgensen AM, Bogeso KP, Peters GH. Molecular dynamics simulations of Na(+)/Cl(-)-dependent neurotransmitter transporters in a membrane-aqueous system. *ChemMedChem* 2007;2:827–840.
46. Celik L, Sinning S, Severinsen K, Hansen CG, Moller MS, Bols M, Wiborg O, Schiott B. Binding of serotonin to the human serotonin transporter. Molecular modeling and experimental validation. *J Am Chem Soc* 2008;130:3853–3865.