bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Deep Mutagenesis of a Transporter for Uptake of a Non-Native Substrate Identifies Conformationally Dynamic Regions

Heather J. Young,[1,] Matthew Chan,[2,]* Balaji Selvam,[2,]* Steven K. Szymanski,[1] Diwakar Shukla,[2,3,4] and Erik Procko,[1,3]*

1 Department of Biochemistry, University of Illinois, Urbana, IL, USA

2 Department of Chemical and Biomolecular Engineering, University of Illinois, Urbana, IL, USA

3 Cancer Center at Illinois and Center for Biophysics and Quantitative Biology, University of Illinois, Urbana, IL, USA

4 National Center for Supercomputing Applications and NIH Center for Macromolecular Modeling and Bioinformatics, University of Illinois, Urbana, IL, USA

* *These authors contributed equally.*

Correspondence: procko@illinois.edu; diwakar@illinois.edu

Title Running Head: Deep mutagenesis of human SERT

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Abstract

The serotonin transporter, SERT, catalyzes serotonin reuptake at the synapse to terminate neurotransmission via an alternating access mechanism, and SERT inhibitors are the most widely prescribed antidepressants. Here, deep mutagenesis is used to determine the effects of nearly all amino acid substitutions on human SERT surface expression and transport of the fluorescent substrate analogue APP+, identifying many mutations that enhance APP+ import. Comprehensive simulations of the entire ion-coupled import process reveal that while binding of the native substrate, serotonin, reduces free energy barriers between conformational states to promote SERT dynamics, the conformational free energy landscape in the presence of APP+ instead resembles Na^+ bound-SERT, with a higher free energy barrier for transitioning to an inward-facing state. The deep mutational scan for SERT-catalyzed import of APP+ finds mutations that promote the necessary conformational changes that would otherwise be facilitated by the native substrate. Indeed, hundreds of gain-of-function mutations for APP+ import are found along the permeation pathway, most notably mutations that favor opening of a solvent-exposed intracellular vestibule. The mutagenesis data support the simulated mechanism in which the neurotransmitter and a symported sodium share a common cytosolic exit pathway to achieve coupling. Furthermore, the mutational landscape for SERT surface trafficking, which likely filters out misfolded sequences, reveals that residues along the permeation pathway are mutationally tolerant, providing plausible evolutionary pathways for changes in transporter properties while maintaining folded structure.

Keywords: deep mutagenesis, transporter, SERT, alternating access mechanism, molecular dynamics, conformational free energy landscape

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

INTRODUCTION

The serotonin transporter (SERT/SLC6A4) catalyzes the reuptake of serotonin from the synapse to terminate synaptic transmission (*1*, *2*). As evidenced by the effects of pharmaceuticals and illicit drugs that inhibit SERT, and genetic studies of knockout mice (*3*) and natural human variants (*4*–*7*), dysregulation of serotonergic signaling is associated with broad psychopathological symptoms, including major depression, anxiety, obsessive compulsive disorder, post-traumatic stress disorder, and attention-deficit/hyperactivity disorder, amongst others (*8*). Similar to other closely related monoamine neurotransmitter transporters for dopamine (DAT/SLC6A3) (*9*, *10*) and norepinephrine (NET/SLC6A2) (*11*), SERT mediated serotonin (5-hydroxytryptamine; 5HT) translocation is coupled to the favorable dissipation of a Na^+ electrochemical gradient. While the stoichiometry for coupled transport remains unclear (*12*), it most likely involves the Cl^--dependent import of one Na^+ and one positively charged 5HT into the cell by SERT (*13*–*15*) using an alternating access mechanism (*16*, *17*), in which the transporter in an outward-facing (OF) conformation captures extracellular substrate, transitions through an occluded (OC) state where the central substrate binding sites become closed and inaccessible to the outside, and then an intracellular exit pathway opens in an inward-facing (IF) conformation for cytosolic release (*18*, *19*). SERT then returns to an outward-facing conformation through the antiport of one K^+, together with possible movement of a proton, for an overall electroneutral cycle (*12*, *13*). Alternative stoichiometries and even channel-like conduction states may occur under different conditions (*15*, *20*–*23*).

Monoamine transporters and other neurotransmitter:sodium symporters (NSS) share a common architecture (*24*) known as the LeuT fold, named after the bacterial LeuT transporter, which was the first member structurally characterized at atomic resolution (*25*). Static crystal

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

structures of human SERT and *Drosophila* DAT have revealed atomic details of OF and OC conformations bound to drugs and substrates (*26–31*). These structures have resolved two Na^+ sites (Na1 and Na2) and the location of Cl^- binding, as well as specific contacts to the neurotransmitter aromatic ring and amine moieties within subsites (*32*) of the central orthosteric (or S1) binding site. In addition, structural, biochemical and pharmacological studies have shown that substrates entering the outward-facing extracellular vestibule engage an allosteric (or S2) site immediately adjacent to critical gating residues, before moving past the open gate to the orthosteric site (*26*, *33–35*). Cryo-EM analysis of detergent-solubilized SERT in the presence of ibogaine, a psychoactive plant product that stabilizes IF conformations, allowed for the characterization of OF, OC and IF-like states at moderate resolution (*19*), showing that regions already implicated in bacterial and mammalian NSS homologues through multiple methods (*18*, *36–41*) move to open a solvent-accessible intracellular vestibule. In particular, transmembrane helix (TM) 1a moves away from the helical bundle. These motions are supported by comprehensive HDX-MS analysis of LeuT (*42*) and SERT (*43*). However, the overall molecular mechanism for ion-coupled neurotransmitter transport remains incompletely understood. How is stoichiometry achieved? How does neurotransmitter move from the orthosteric site to the cytosol? How do specific residue-substrate interactions help orchestrate the conformational changes associated with an alternating access mechanism? Even the most recent cryo-EM structures of ibogaine-bound SERT only confirmed the general region for neurotransmitter exit, and the authors were uncertain as to the exact exit pathway amongst multiple possibilities (*19*). Knowing the mechanism in atomic detail can assist in understanding how genetic variants are associated with disease, and may aid the development of drugs targeting discrete conformational states.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

We recently described molecular dynamics (MD) simulations of the entire SERT-catalyzed import process of 5HT (*44*), accomplished using adaptive sampling to efficiently explore conformational space. The calculated highest flux pathway for substrate binding events and conformational changes lead to several important conclusions, including the identification of key residues involved in substrate recognition at the extracellular vestibule and the prediction that the neurotransmitter exits along a pathway beginning at the Na2 site that is surrounded by TM1a, TM5, TM6b and TM8.

Here, we use deep mutational scanning (*45*) to define the mutational landscape of SERT for surface expression, which indicates how mutations effect folding and escape from intracellular protein quality control (*46*), and substrate transport. Deep mutagenesis involves (i) creating a library of protein variants, (ii) using in vitro selection to enrich for the variants with high activity, and (iii) measuring the enrichment or depletion of all the variants using next generation sequencing. Deep mutational scanning has proven insightful for indirectly interrogating protein conformational states in living cells (*46–49*), and datasets that include higher order mutations can even assist in modeling small protein structures with atomic resolution (*50*, *51*). We build on this body of work by now applying deep mutagenesis to a human transporter that undergoes substantial conformational change. By using a non-native substrate that is unable to facilitate isomerization to an inwards-facing conformation, we find extensive gain-of-function mutations in SERT that provide insight on dynamic regions in close agreement with simulation.

RESULTS

Mutational landscapes of SERT for plasma membrane localization and substrate import

We sought to understand how SERT sequence relates to conformational dynamics and mechanism through deep mutagenesis, which demands an effective selection strategy to

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

distinguish low and high activity SERT sequence variants. This is not possible using transport assays with radioactive 5HT. Instead, *in vitro* selection was based on the uptake of the fluorescent monoamine neurotransmitter analogue APP+ (Figure S1), permitting cells expressing active and inactive SERT variants to be readily distinguished and separated by fluorescence. We used a fluorescence-based transport assay to assess SERT activity by flow cytometry, which could then be seamlessly integrated with fluorescence-activated cell sorting (FACS) to screen libraries of SERT variants. A synthetic gene encoding human SERT with a c-myc epitope tag (flanked by gly/ser-rich linkers) inserted between residues Asp216 and Asn217 of EL2 was constructed; the homologous site in *Drosophila* DAT features a large domain insertion. When expressed from an epsiomal plasmid in Expi293F cells (a suspension culture derivative of human HEK293 that has advantages for FACS-based selections of large libraries), surface localized transporter was detected with a fluorescent anti-myc antibody, while uptake of APP+ was simultaneously observed (Figures S2A and S2B). Myc-tagged SERT has ion dependence consistent with untagged human SERT (*14*, *52–54*) (Figure 6E), appropriate neurotransmitter selectivity (Figure 6D), and exhibits a similar K_M for APP+ as for the endogenous substrate 5HT (*53*, *54*). However, transport velocity of myc-tagged SERT is reduced 4-fold compared to the native protein, while insertion of a shorter gly/ser-rich sequence at the same site had no significant effect (Figure S2C). We speculate that the bulkier myc tag may partially obstruct substrate access to the extracellular vestibule, a limitation that we considered offset by the capability of measuring transport and surface localization simultaneously. Furthermore, a subset of mutations with enhanced APP+ transport discovered by deep mutagenesis were confirmed to also enhance activity in native untagged SERT (Figure S2D), and the presence of the myc tag is therefore very unlikely to have impacted scientific conclusions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Tagged SERT was mutagenized to generate nearly all single amino acid substitutions. To increase sampling during FACS-based selection, substitutions were split across two libraries spanning the N- (residues 2-325) and C- (residues 326-630) termini, each containing ~6,000 mutants. The libraries were diluted with carrier DNA encoding a viral factor for enhanced replication of the episomal plasmid (*46*) and transfected into Expi293F cells. A large excess of carrier DNA ensures that typically only one SERT coding variant is expressed in any cell (*48*), providing a tight link between genotype and phenotype. Cell libraries were incubated with APP+ at a concentration below the K_M and stained with anti-myc antibody conjugated to an orthogonal fluorochrome to simultaneously assess surface expression. Cells expressing SERT at the plasma membrane, or expressing SERT and transporting the fluorescent substrate, were collected via FACS (Figures S2E and S2F). After deep sequencing, mutation enrichment ratios were calculated by comparing variant frequencies from RNA transcripts in the sorted cells to the naive DNA library. These data experimentally define the mutational landscapes for SERT surface expression and APP+ transport (Figure 1).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

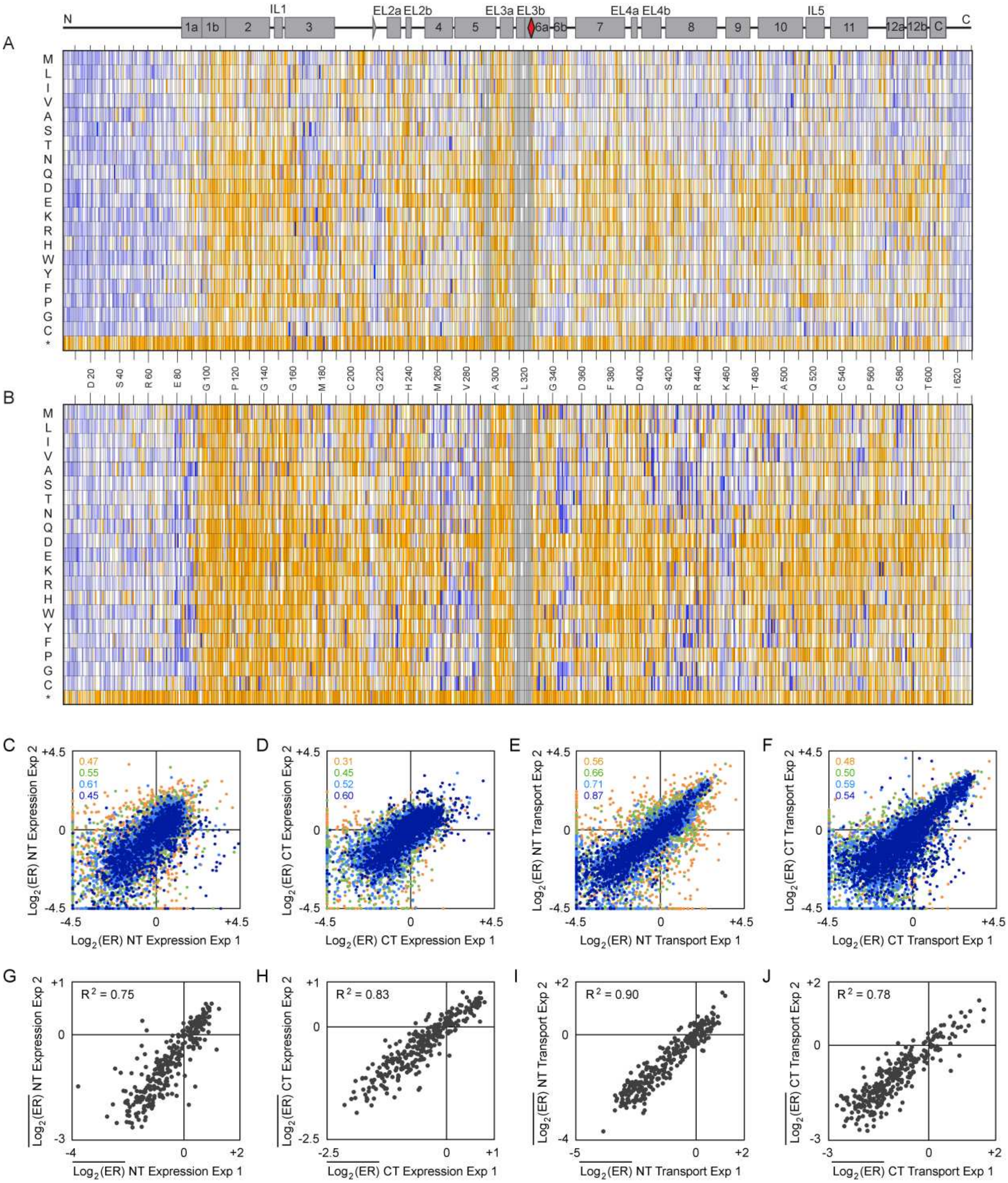

Figure 1. SERT mutational landscapes for surface expression and APP+ transport. **(A, B)** SSM libraries of SERT spanning N- and C-terminal residues (the red diamond in the secondary structure schematic at top indicates the dividing point between the libraries) were sorted for (A) surface expression and (B) cell uptake of APP+. Log_2 enrichment ratios are plotted from ≤ -3 (depleted mutations, orange) to 0 (neutral, white) to ≥ +4 (enriched, dark blue). Data are

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

averaged from two independent experiments. Missing mutations (< 10 reads in the naïve libraries) are grey. *, stop codons.

(C-F) Agreement between $\log_2$ enrichment ratios from replicated selections for surface expression of the SERT (C) N-terminal and (D) C-terminal SSM libraries, or for APP+ transport of SERT (E) N-terminal and (F) C-terminal SSM libraries. Agreement between independent experiments generally increases for more frequent sequence variants (orange, 10-99 reads in the naïve library; green, 100-199 reads; light blue, 200-399 reads; dark blue, 400+ reads). R^2 values are correspondingly colored in the upper-left corners.

(G-J) Agreement between conservation scores (defined as the mean of the $\log_2$ enrichment ratios at a single residue position) from replicated selections for surface expression of SERT (G) N-terminal and (H) C-terminal SSM libraries, or APP+ transport of (I) N-terminal and (J) C-terminal libraries.

Enrichment ratios are correlated between independently replicated selection experiments (Figures 1C-1F). Agreement between replicate experiments tends to increase for highly abundant sequence variants, presumably due to increased sampling. Metrics for mutational tolerance such as sequence Shannon entropy do not distinguish between positions that tolerate many neutral mutations versus positions that are hot spots for gain-of-function mutations. We therefore calculate a simple, unambiguous average of the $\log_2$ enrichment ratios for all 20 amino acid possibilities at each position, which we call a residue conservation score. Conservation scores are highly correlated between replicates (Figures 1G-1J) and define regions of sequence that are tightly conserved for function, or alternatively are under positive selection to change.

Mutational scans of transmembrane proteins for plasma membrane localization indicate whether mutations adversely impact folding and structural stability, as folded proteins are anticipated to preferentially escape intracellular quality control machinery (*46*). In the mutational landscapes for both SERT surface expression and APP+ transport, mutations within secondary structural elements are generally deleterious, in particular polar mutations within transmembrane segments (Figures 1A and 1B). In mutational scans of G protein-coupled receptors (*48*), proline substitutions, which are disruptive to helical conformations, were depleted in regions (including

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

termini) that have ordered structure, yet no such depletion of proline substitutions is observed in the cytosolic termini of SERT. Instead, the SERT cytosolic tails tolerate most mutations, strongly indicating that these regions lack critical folded structure despite previous modeling to the contrary (*55*). In comparison, mapping conservation scores for each residue position to the crystal structure of SERT (*26*, *27*) highlights regions of folded structure under sequence constraints (Figure 2). The selection of SERT mutants for surface expression reveals conservation within packed hydrophobic cores ‘above’ and ‘below’ the central orthosteric binding site (Figure 2B), while the protein sequence along the permeation pathway is tolerant of many substitutions. This suggests mutations in the permeation pathway, which may be desirable for the evolution of new transporter specificities and coupled ion stoichiometries, may frequently occur without major disruption of the basic protein architecture and fold. In the deep mutational scan for APP+ import, sequence conservation now extends towards residues lining the entrance vestibule and into the neurotransmitter and ion binding sites (Figure 2C).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

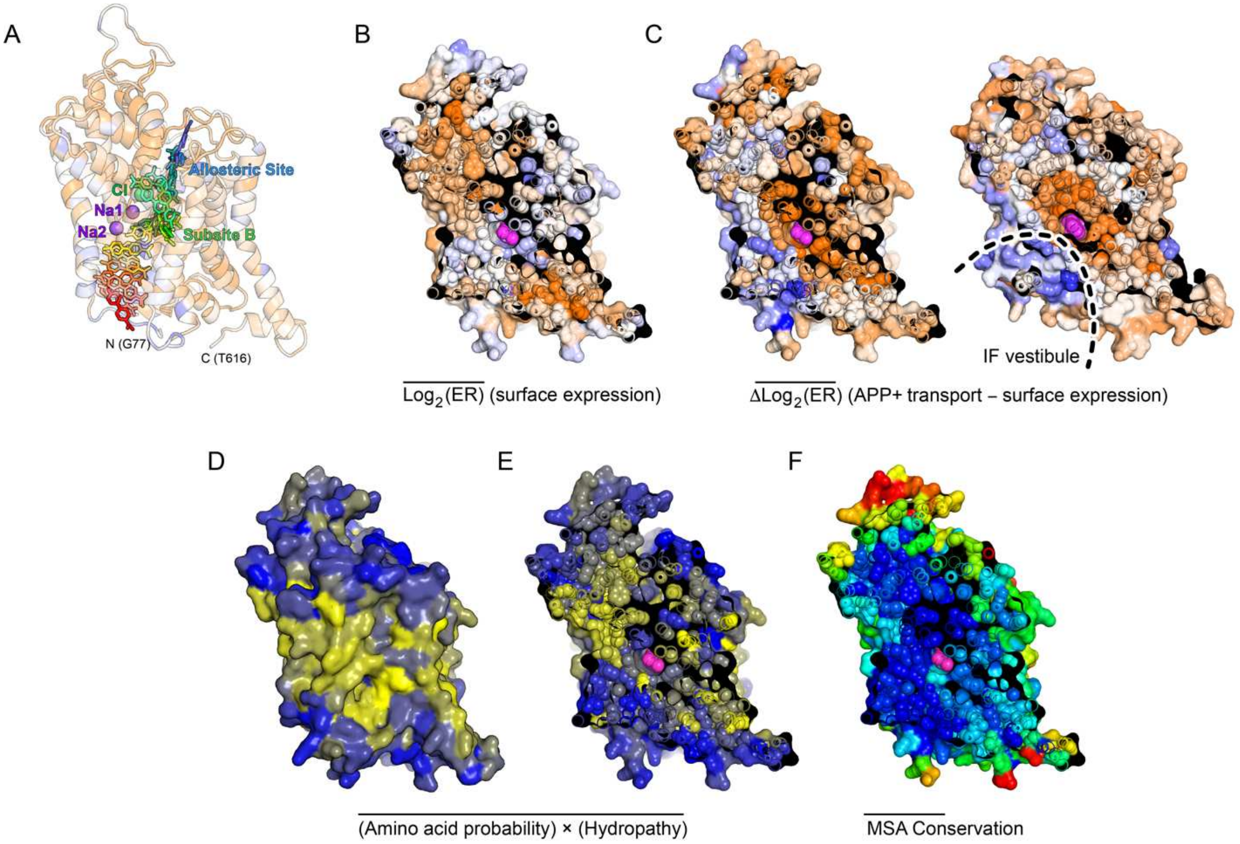

Figure 2. SERT mutations that increase APP+ import are enriched in the intracellular exit pathway.

(A) Overlaid MD snapshots of APP+ translocation, from when APP+ enters the extracellular vestibule (blue) to its cytosolic exit (red). The positions of ions in the OC state are shown with spheres. SERT is shown as a semi-transparent ribbon and colored by experimental sequence conservation for APP+ transport, showing regions where mutations are depleted in orange to regions where mutations are enriched in dark blue. Also see structure file in Supporting Information online.

(B) A cross-section through SERT in a partial OF state (PDB 5I73; bound to (*S*)-citalopram in magenta) colored by conservation for folding and surface expression. Tightly conserved residues in the core are orange.

(C) Cross-sections through SERT in a partial OF state at left and in an ibogaine-bound IF-like state at right (PDB 6DZZ), highlighting in orange residues within the substrate-binding sites and extracellular cavity that become conserved for APP+ transport, while residues that are hot spots for gain-of-transport mutations are dark blue and map to the intracellular exit pathway. Drugs are magenta.

(D, E) Enrichment ratios for APP+ transport were converted to amino acid probabilities, weighted by hydropathy (yellow, hydrophobic; blue, polar), and averages are mapped onto the **(D)** surface and **(E)** core of SERT (PDB 5I73). Residues lining the entrance cavity and putative exit pathway tolerate polar substitutions.

(F) Natural conservation based on a sequence alignment of 150 SERT homologs (blue, conserved; to red, variable). Core, entrance, and exit pathways are conserved in natural history.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

The mutational scan identifies hundreds of gain-of-function mutations based on positive enrichment after sorting for APP+ import. We further tested by targeted mutagenesis 22 predicted gain-of-function mutations that are spatially dispersed across the structure and validated 21 as trending toward increased APP+ uptake, with 11 having statistical significance (Figure S3). Gain-of-function phenotypes persist when the mutations are introduced into the untagged native cDNA (Figure S2D). We also tested previously reported allelic variants G56A, I425V, and K605N, which have increased 5HT transport due to altered PKG signaling (*6*, *7*, *56*, *57*), and found these did not have highly elevated APP+ uptake when expressed individually, consistent with their lack of positive enrichment in the mutational scan (Figures S2D and S3A). Cyclic GMP-dependent signaling is therefore unlikely to be an important pathway for modulating SERT activity in the experimental system. Expression of the mutants was generally similar to or less than wild type and cannot explain the large increases in transport activity (Figures S3B and S3C). Targeted mutagenesis therefore validates predictions from the mutational scan that many amino acid substitutions increase APP+ transport kinetics.

Conformational transition to the inward-facing state is rate-limiting for APP+ import

Strikingly, the bulk of mutations that enhance APP+ uptake are clustered in the intracellular half of the transporter (Figure 2C). This clustering of gain-of-function mutations is most noticeable at TM1a, TM6b, and the intracellular segment of TM8, loosely focused near the axis of pseudosymmetry and tightly aligned to the simulated neurotransmitter exit pathway (*44*). After weighting the enrichment ratios by hydropathy and averaging, we find that this region of SERT is highly tolerant of diverse mutations including polar substitutions (Figures 2D and 2E). Mutations in this region will disrupt buried interactions along the exit pathway and thereby

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

facilitate opening of a solvent-exposed intracellular vestibule. That these mutations enhance APP+ uptake is exactly what one would expect if formation of the IF state is rate-limiting.

As previously reported (*44*), the entire Na$^+$-coupled import process of the physiological substrate 5HT has been simulated using Markov state model (MSM) based adaptive sampling (*58*–*60*) to efficiently explore conformational space. MD with adaptive sampling has been successfully used to simulate complex transport processes in other systems (*59*, *61*). MSM-weighted SERT simulation data are projected onto a coordinate system defined by distances between extracellular and intracellular gating residues (*38*, *44*), thereby visually displaying the relative free energies and barriers between kinetically relevant intermediate states (Figure 3). SERT free of neurotransmitter substrate (i.e. Na$^+$-SERT) readily transitions between OF and OC states (Figure 3A) and only in the presence of the cognate substrate 5HT does the IF state become substantially occupied (Figure 3B), with a reduction in the free energy barrier for the OC-IF transition. Here, the simulations are replicated, but with 5HT replaced by APP+ in the simulation box as the substrate. Na$^+$-bound SERT in an OF conformation obtained from Na$^+$-SERT simulations (*44*) was used to seed MD simulations of APP+ import totaling ~370 μs (referred to as APP-SERT simulations; Figures 3C and S4).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

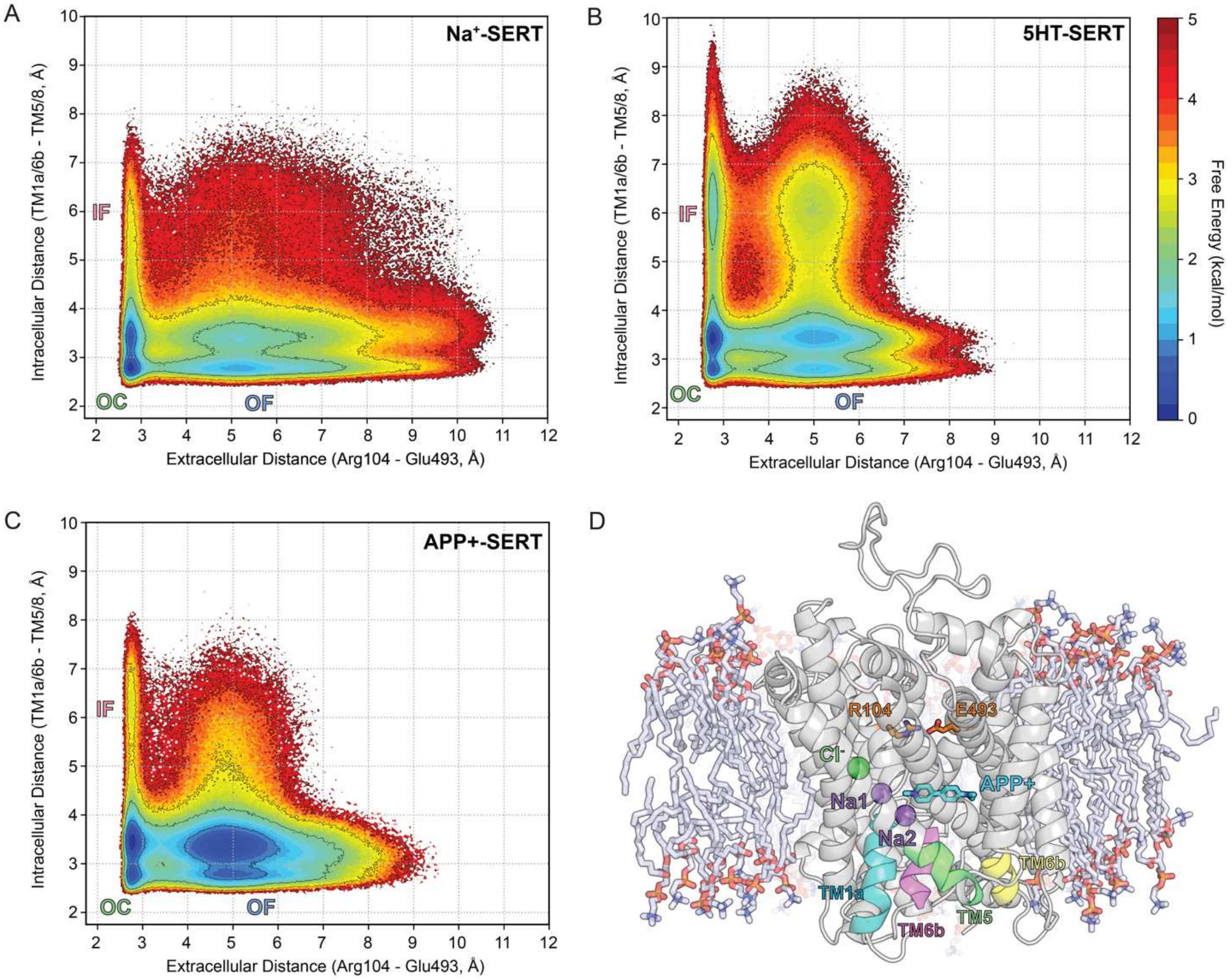

Figure 3. Simulated conformational free energy landscapes of SERT.

(A, B) Relative free energies from MSM-weighted simulation data plotted against the distances between extracellular and intracellular gating residues for (A) Na$^+$-SERT and (B) 5HT-SERT. These simulation data were previously reported (*44*) and are reproduced here for comparison. The outward-facing (OF) crystal structure (PDB 5I73) was used as the starting structure for MD simulation and transitioned to occluded (OC) and inward-facing (IF) states.

(C) Calculated free energy landscape of SERT during the APP+ translocation process.

(D) MD snapshot of membrane-embedded SERT with APP+ bound in subsite B and ions bound in respective sites. Extracellular gating residues Arg104 and Glu493 shown in orange sticks. Intracellular distance was measured based on distances between TM1a, TM6b and the cytoplasmic base of TM5 and TM8, colored in teal, magenta, green, and yellow, respectively.

The simulations reveal that the free energy basin corresponding to IF conformations remains at a higher free energy value in the presence of APP+ (Figure 3C), resembling the conformational free energy landscape for Na$^+$-SERT (Figure 3A). The free energy barriers for OC-

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

IF transitions in Na^+-SERT and APP-SERT are calculated to be 3.43 ± 0.06 kcal/mol and 2.89 ± 0.08 kcal/mol, respectively, which are higher than the estimated 1.66 ± 0.04 kcal/mol barrier for the OC to IF transition when 5HT is bound and permeating along the transport pathway. Indeed, formation of the IF state is predicted to be rate-limiting for APP+ import, whereas 5HT interactions facilitate dynamic motions necessary for transport by reducing free energy barriers and stabilizing the IF state (compare Figures 3B and 3C). Similar conclusions regarding differences between substrates have been drawn from experimental observations of bacterial LeuT; only a good substrate such as alanine readily facilitates transition from the OC to IF state (*38*, *62*). The conformational free energy landscape of APP+ import also shows OF states to be further stabilized as the bulkier substrate occupies the extracellular vestibule and the allosteric site. Partial OF-IF-like conformations are also observed, which we previously referred to as an hourglass (HG) state where both gates are open but there is a central constriction (*44*). We speculate that HG conformations may be associated with SERT conduction states (*15*, *20–23*).

Natural evolution can be considered akin to a deep mutagenesis experiment for protein function as it relates to organism fitness, and we note that residues of SLC6 family members along the permeation pathway are conserved between species (Figure 2F). Furthermore, an exome database shows little human variation (*63*). SERT has therefore reached an optimum (or close to optimum) sequence for its physiological function of ion-coupled 5HT transport within the constraints of natural evolution. In contrast, our deep mutational scan identifies mutations that increase import of the unnatural substrate APP+ for which the protein sequence is suboptimal. In essence, mutations are found that impose on the APP-SERT conformational free energy landscape the necessary changes that would ordinarily be facilitated by interactions with the native substrate 5HT.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Based on crystal structures and computational modeling, the N-terminal cytosolic tail of LeuT, in particular LeuT residue Arg5 (equivalent to Arg79 of SERT), forms an interaction network that stabilizes a closed conformation of the cytoplasmic exit cavity in the OF and OC states (*25*, *38*, *41*). The simulations show dynamic motions of the equivalent residues in SERT (a.a. 78-82), with Arg79 forming transient salt bridges to Glu444 and Asp452 to restrict opening of the cytoplasmic exit. Consistent with transitions to the IF state being rate-limiting, mutations to Arg79 and surrounding cytoplasmic gating residues increase APP+ import by destabilizing closure of the intracellular vestibule. We note that alanine substitution of Arg79 (*64*) is deleterious for SERT-catalyzed transport of 5HT and likely imbalances the conformational equilibrium. (Likewise, mutations to Asp437 in the Na2 site, which are discussed further below, also increase APP+ import but reduce 5HT transport (*52*).) It is only for APP+, which is unable to facilitate appropriate conformational dynamics, that these mutations impart a gain-of-function phenotype, providing a unique mutational data set that functionally maps the substrate exit pathway.

Deep mutagenesis supports the simulated binding mechanism of Na$^+$ and neurotransmitter

The highest flux pathway for ion/APP+ binding events and conformational changes was determined by transition path theory (*65*) and predicts an ordered process for APP+ import identical to that previously simulated for 5HT (Figure S4). Na$^+$ first binds an *apo*-SERT at the Na1 site, followed by a second Na$^+$ occupying the Na2 site. APP+ then enters the transporter through the extracellular vestibule and moves to the orthosteric site, with localized dynamics around extracellular loop 4 (EL4) and the extracellular gate, consistent with static structures (*18*, *19*), HDX data (*42*, *43*), and voltage-clamp fluorometry (*66*). Unlike the endogenous substrate 5HT, the increased structural rigidity of APP+ (Figure S1) results in its translocation ‘down’ the extracellular vestibule to be more constrained, with decreased rotational freedom. In the mutational

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

landscape, hydrophobic substitutions are tolerated or weakly enriched at positions Phe335 and Glu494, where they pack against the APP+ aromatic ring moieties (Figure 4A).

Cl$^-$, accompanied by an extracellular Na$^+$ ion, then engages the extracellular gating residues Arg104 and Glu493, disrupting their electrostatic interactions to allow Cl$^-$ to enter the orthosteric cavity. The additional Na$^+$ ion does not further progress towards the orthosteric cavity and diffuses back to the extracellular space. Once the neurotransmitter and ions are all bound to their respective sites, the extracellular gates close to form OC conformations. The stability of the OC state promotes the solvation of the intracellular vestibule, thereby weakening the hydrogen bonding network involving various intracellular gating residues (Arg79-Glu452, Glu80-Lys275, Asp87-Trp282, Tyr350-Glu444). Opening of the intracellular permeation pathway is associated with displacement of Na$^+$ from the Na2 site into the cytoplasm (consistent with extensive experimental evidence (*36*, *40*, *67*)), followed by cytosolic exit of APP+.

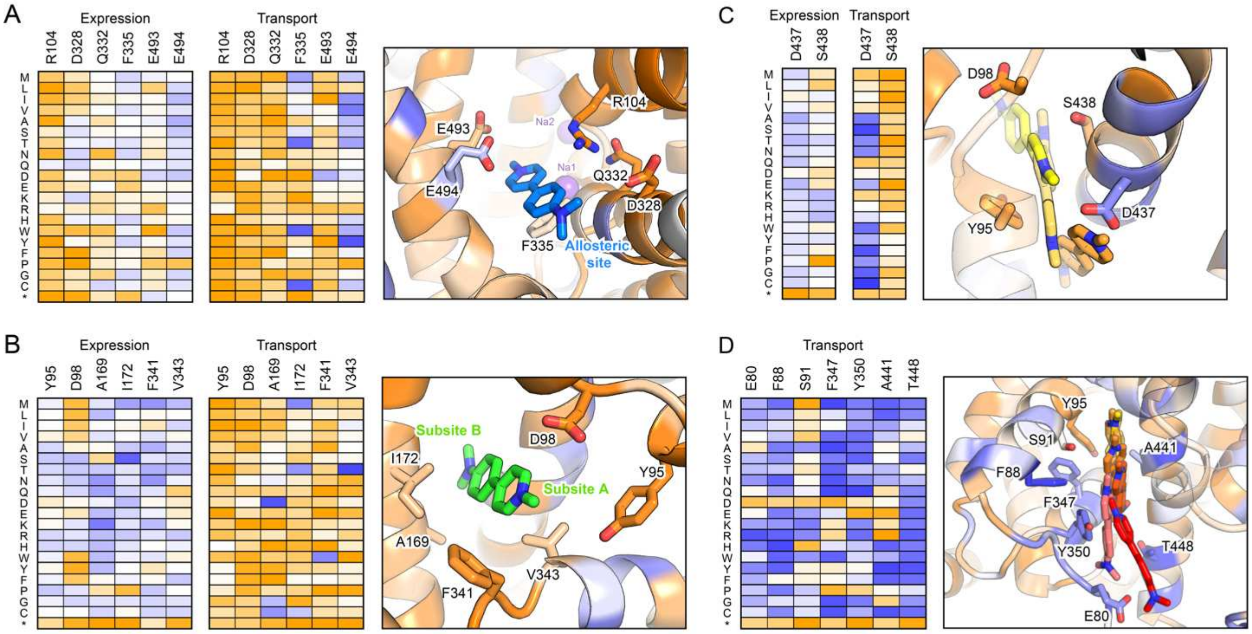

Figure 4. The simulated mechanism for SERT-catalyzed substrate import is supported by the mutational landscape.

Heat maps (colored as in Figure 1) show depleted mutations for SERT surface expression (left) or

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

APP+ transport (right) in orange and enriched mutations in dark blue. Accompanying MD snapshots show SERT as a cartoon representation with residues colored according to conservation score for transport (as described in Figure 2A).
(A) APP+ entering the transporter via the allosteric site.
(B) The neurotransmitter substrate bound to the orthosteric site, straddling subsites A and B.
(**C**) APP+ exiting the orthosteric site through the Na2 site.
(**D**) APP+ permeation along the exit pathway.

Simulations show contacts made by the neurotransmitter to specific subsites (*32*) within the orthosteric binding pocket change as import progresses. The aromatic ring of 5HT shifts from subsite C to B, while flexibility in the aliphatic chain accommodates the amine moiety remaining bound to subsite A (*44*). APP+ never fully straddles subsites A to C, most likely due to steric limitations from having a rigid structure, yet it does interact with subsite B in a binding mode supported by crystallography of dopamine-bound DAT (*30*) and biochemical analyses of inhibitor binding (*32*, *68*, *69*). While residues surrounding the subsites are generally conserved for APP+ import, A169D within subsite B enhances transport (Figure 4B). This mutation is known to also increase the potency of inhibitors (*68*). A169D introduces a carboxylate ~12 Å distal to where the substrate's amine group is positioned at subsite A. This spacing is ideal for accommodating the length of APP+, which through resonance has positive charge distributed on both ends for favorable electrostatics in the A169D mutant. This provides further evidence that monoamine substrates are indeed accommodated in subsite B during the transport process.

In the deep mutational scan, mutations to residues coordinating ions in the Cl^- and Na1 binding sites are highly deleterious for transport activity. However, gain-of-function mutations that enhance APP+ uptake begin with Asp437 in the Na2 site and continue along the entire exit pathway to the cytoplasm (Figures 4C and 4D). There is a near perfect alignment of gain-of-function mutations for APP+ import on TM1a, TM5, TM6b, and TM8 with the simulated

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

permeation trajectory (Figure 4D and structure visualization file in Supporting Information). The Na2 site therefore forms the apex for the cluster of the gain-of-function mutations that delineate the intracellular vestibule. The data are consistent with the simulated release of Na^+ from the Na2 site, followed by the movement of 5HT or APP+ into the exit pathway by first passing through the vacated Na2 site. The release of neurotransmitter from the orthosteric site requires a rotameric shift of Tyr95 that is coupled to Na^+ dissociation at the Na2 site. Tyr95 forms a cation-pi interaction with the substrate amine group and is highly conserved in the mutational scan (Figure 4B). The sharing of a common exit pathway by neurotransmitter and Na^+ beginning at the Na2 site elegantly explains the 1:1 sodium to neurotransmitter stoichiometry of the transport cycle.

Gain-of-function mutations for APP+ uptake at positions outside the permeation pathway are predicted to enhance SERT conformational dynamics.

We clustered the SERT mutational data for surface expression and APP+ transport activity using an unsupervised learning algorithm, Uniform Manifold Approximation and Projection (UMAP) (*70*). UMAP presents an approach to effectively reduce the high dimensional mutational landscape and systematically derive quantitative relationships between a residue's mutational response and its biophysical and biochemical properties (*71*, *72*). Visualization of the mutational landscape on the two-dimensional UMAP space shows that residues cluster based on various biochemical and biophysical properties (Figures 5 and S5). Projection of the mutational landscape for APP+ transport shows that residues naturally cluster based on residue site location (Figure 5B) and hydrophobicity (Figure 5C), once again illustrating the mutational tolerance of the SERT intracellular domain on expression and transport. When compared to biophysical properties obtained from MD simulations, we observed that a residue's mutational activity is related to its solvent accessible surface area (SASA) (Figure 5D) and its dynamic fluctuations (Figure 5E).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Residues that enhance transport activity are associated with larger fluctuations, while residues with low mutational tolerance exhibit low SASA and RMSF, suggesting these residues are buried in the protein core and are important for stabilizing contacts to maintain proper folding. Furthermore, the deep mutational landscape correlates with mutational effects predicted by EVmutation (*73*), underscoring that sequence constraints necessary for proper transporter function are at least partially shared between the in vitro selection and natural evolution.

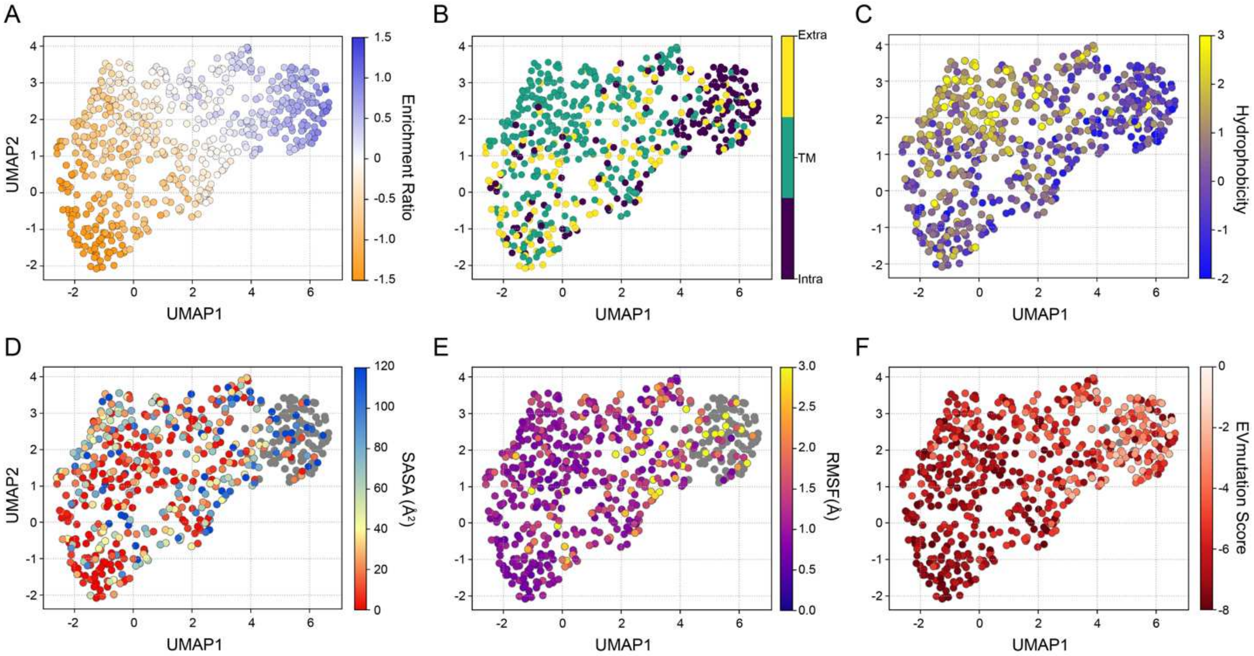

Figure 5. UMAP projection of the mutational landscape for SERT-catalyzed APP+ transport.

Each residue was clustered using the UMAP learning algorithm based on the activities of all amino acid substitutions at that position, and are colored based on different biochemical and biophysical properties: (**A**) APP+ uptake enrichment ratio, (**B**) residue location, (**C**) Kyte-Doolittle hydrophobicity, (**D**) solvent accessible surface area (SASA), (**E**) root mean square fluctuation (RMSF), and (**F**) EVmutation score. Residues 2-75, where SASA and RMSF cannot be calculated, are colored gray in panels (**E**) and (**F**). The first UMAP dimension strongly correlates with enrichment ratio (transport ρ = 0.9015). Projection of the SERT expression deep mutational landscape in 2D UMAP space is available in Supplemental Information Figure S5.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

While less abundant, some mutations outside the simulated permeation pathway are also enriched for APP+ import, including mutations at helical hinges that flex during structural transitions (e.g. A173G, P499K, S585P), or on neighboring helices immediately adjacent (e.g. F263C, P499K, G582K); these substitutions may stabilize unwound structural features within the membrane phase, facilitate dynamic motion, or stabilize conformational states. The mutational tolerance of Ala173 at a helix kink is particularly striking; nearly all substitutions are deleterious except for the small amino acid glycine that can increase flexibility (Figure 6B). DAT and NET both have glycine at this position. We also find a cluster of gain-of-function mutations for APP+ import within the hydrophobic core where TM1b packs against EL4 (Figure 6C); substantial EL4 motion in bacterial transporters can occlude the extracellular vestibule (*74*), while more limited EL4/TM1b motions are observed in cryo-EM structures of ibogaine-bound SERT (*19*) and in the simulations. In general, mutations in this region that increase APP+ import substitute large hydrophobics for smaller side chains that reduce packing and are therefore anticipated to increase dynamics, possibly creating localized molten globule-like behavior.

We further investigated four SERT mutations that increase APP+ import, chosen to represent different structural regions: R79V in the cytoplasmic gating loop (Figure 6A), F88H on TM1a proximal to the exit pathway (Figure 6A), A173G at a helix kink in TM3 (Figure 6B), and L406I in EL4b where it packs against TM1b (Figure 6C). All mutants retained selectivity for 5HT over dopamine and norepinephrine (Figure 6D), consistent with these mutations impacting kinetic barriers or conformational equilibria rather than binding interactions within the orthosteric site. Na^+ and Cl^- dependency of APP+ import is partially lost in two of the mutants, R79V and F88H (Figure 6E), which both lie on the exit pathway and are predicted to destabilize closure of the intracellular vestibule. Correct ion coupling for symport therefore demands appropriate free energy

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

differences between conformational states, ensuring that all ions and substrate must bind to energize structural transitions.

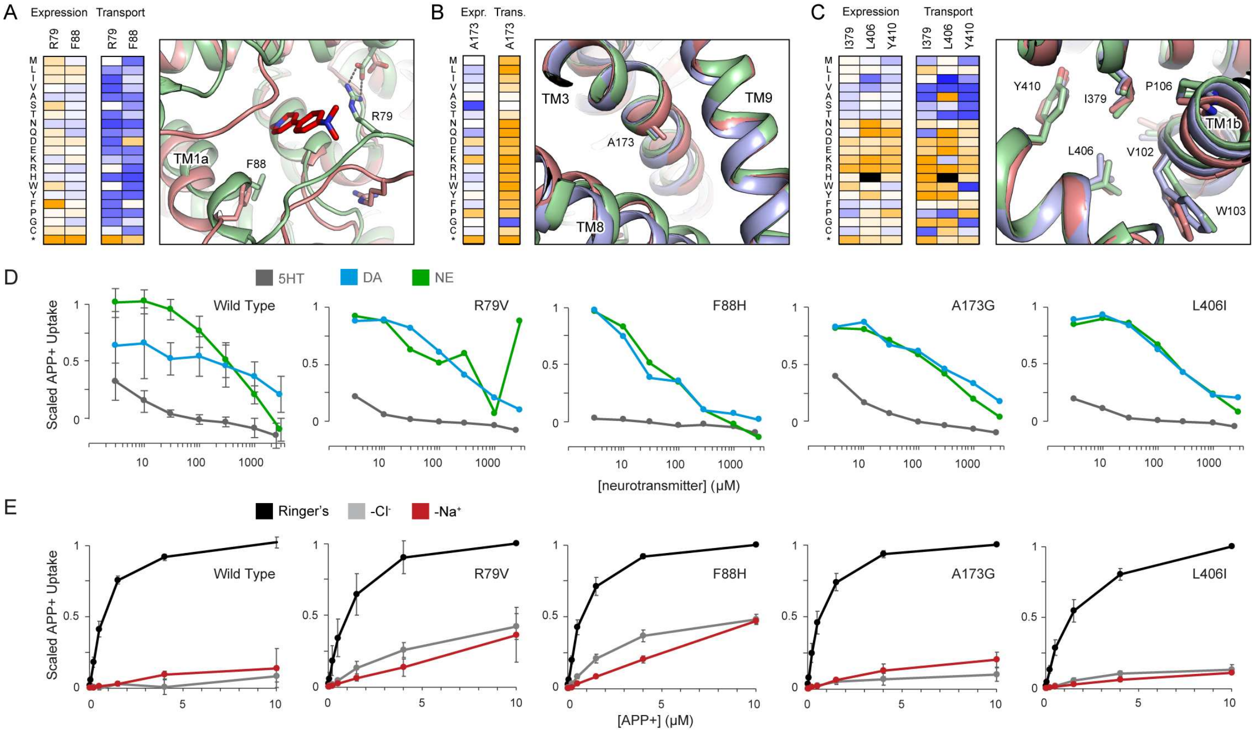

Figure 6. Exit pathway mutations can cause a partial loss of Na$^+$ and Cl$^-$ dependency.
(A-C) Superimposed MD snapshots (blue, OF state; green, OC state; and pale red, IF state) showing dynamic regions. Experimental enrichment ratios for select residues are plotted alongside, colored orange (deleterious mutations) to dark blue (enriched mutations) as in Figure 1.
(A) Motion around the intracellular gate.
(B) A hinge in TM3.
(C) The hydrophobic core below EL4, adjacent to the extracellular gate.
(D) APP+ uptake in SERT-expressing cells was more potently inhibited by 5HT (grey) than by norepinephrine (NE, green) or dopamine (DA, blue). (Mean; WT n = 5, R79V n =2, F88H n =2, A173G n = 3, L406I n = 2. For WT, error bars show SD.)
(E) APP+ uptake by SERT-expressing cells was measured in Ringer's solution (black), with replacement of Na$^+$ (red) or Cl$^-$ (grey) ions. Transport was Na$^+$ and Cl$^-$ dependent for WT, A173G, and L406I SERT, while showing a partial loss of ion dependence in R79V and F88H mutants. (WT n = 4, R79V n = 4, F88H n = 3, A173G n = 4, L406I n = 3, mean ± SD)

DISCUSSION

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Protein sequences are highly optimized over natural history for their physiological functions. Here, SERT variants were selected for enhanced transport of an unnatural substrate for which the protein sequence is sub-optimal, and which simulations indicate is unable to promote the conformational changes that drive transport. As a consequence, hundreds of gain-of-function mutations were discovered that richly inform molecular mechanism and provide insight into conformational changes that limit transport kinetics. We conclude that in the absence of the cognate substrate, there is a high free energy barrier for transition to the IF state, which will maintain ion-neurotransmitter coupling and prevent ion leakage, and the symported Na^+ and neurotransmitter share a common exit pathway that begins with the Na2 site at the apex of the intracellular vestibule. This proposed import mechanism is likely shared with related neurotransmitter:sodium symporters.

The technical requirements for *in vitro* selection restricted investigations to the fluorescent neurotransmitter analogue APP+ (*75*), which is larger than 5HT with reduced configurational entropy (Figure S1), and its import was therefore sensitive to mutations along the permeation pathway. The concept of discovering mechanism by studying an alternative substrate is widely practiced in the field of enzyme directed evolution (*76*), and engineering of major facilitator superfamily transporters to redirect substrate preference has illuminated important motifs and sequence features (*77–80*). Notably amongst studies of NSS homologues, rational mutagenesis switched LeuT from being inhibited by to now transporting tryptophan, demonstrating that substrate complementarity is necessary for driving the OF-OC structural transition (*81*). Our work here similarly uses an unnatural substrate to inform the mechanism of monoamine NSS transporters through the discovery of gain-of-function mutations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

SERT self-associates into poorly defined oligomers containing up to 8 subunits (*82*, *83*), consistent with aggregation or lipid microdomain localization as opposed to distinct protein-protein interactions. The structures of human SERT and *Drosophila* DAT are both monomeric (*19*, *26*, *30*), and while detergent extraction can dissociate protein-protein interfaces, it nonetheless suggests against a specific, tight oligomer. Crystal structures of LeuT obtained in a dimer suggested that the two monomers may form an interface at TM12 (*18*). Periole et al. conducted coarse-grained SERT simulations and observed the probability of various dimer interfaces involving different transmembrane helices, of which symmetric TM12-TM12 and asymmetric TM7-TM12 dimers were most favorable (*84*). In light of these studies, it remains unknown whether SERT forms specific oligomers that are biologically relevant. In investigations of a dimeric G protein-coupled receptor that is closely related to metabotropic neurotransmitter receptors, deep mutagenesis identified membrane-exposed surface patches that were relatively highly conserved during in vitro selection; the data guided the integrative structure determination for the protein's dimeric architecture that closely agreed with cryo-electron microscopy (*46*). Searching for surface patches on SERT that are conserved in the deep mutational scans, the landscapes for SERT expression and APP+ transport identify the membrane-exposed surface formed by TM2 and TM7 as a candidate oligomerization or lipid-interaction site. However, these helices also form critical contacts to Cl^- and Na^+ within the protein core, and their higher conservation may simply reflect their importance for proper folding with bound ions. The mechanisms by which SERT proteins self-associate requires further investigation.

The mechanistic details described here may provide insights into how SERT kinetics are regulated *in vivo*. For example, syntaxin 1A interacts with the SERT N-terminus to dampen excess Na^+ flux during 5HT transport (*23*, *85*), and it may be that syntaxin 1A stabilizes the OF state with

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

a bound N-terminal gating loop to ensure free energy barriers are sufficiently high for strict coupled transport. By comparison, phosphorylation of Thr276 by PKG when SERT is in the IF conformation increases transport velocity, possibly by preferentially stabilizing the IF state (*57*). Hence under certain physiological conditions, nature may have exploited features of the conformational free energy landscape to increase kinetics or change ion coupling and flux.

METHODS

Cell Culture. Expi293F cells (ThermoFisher) were cultured in Expi293F expression medium (ThermoFisher) at 37ºC, 8% CO_2, 125 rpm.

Reagents. APP+ was from Aobious, prepared as a 100 mM solution in DMSO, and stored at -80°C. Neurotransmitters 5HT, norepinephrine, and dopamine were obtained from Sigma-Aldrich and prepared as 100 mM (dopamine and norepinephrine) or 25 mM (5HT) solutions in water. Neurotransmitter stocks were stored at -80°C until use.

Plasmid Construction. Codon optimized human SERT was assembled from synthetic DNA fragments (IDT) and inserted into the NheI-XhoI sites of pCEP4 (Invitrogen). The use of a synthetic gene ensures there are no sequence features, such as GC-rich tracts, resistant to PCR-based mutagenesis, and therefore aids in library construction. A c-myc epitope tag flanked by a glycine / serine rich linker was inserted into EL2 between residues 216 and 217. A version without the tag was created by PCR-based mutagenesis. For protein expression from the native human cDNA, SERT was subcloned from pcDNA3-hSERT (Addgene 15483)[68] into Kpn1-Xho1 sites of pCEP4. A c-myc epitope tag was inserted into EL2 using overlap extension PCR. All constructs featured a consensus Kozak sequence. Important plasmids are deposited with Addgene.

Library Construction. Two separate SSM libraries were created from plasmid pCEP4 encoding

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

codon optimized myc-SERT. The SERT N-terminal library covered amino acids 2-325, and the C-terminal library covered amino acids 326-630. For each library, overlap extension PCR using primers containing degenerate NNK codons was applied (*86*), and the pooled PCR products were cut and ligated into NheI-NotI or HindIII-XhoI sites for the N-terminal and C-terminal libraries, respectively. Ligation products were electroporated into NEB5α cells (New England Biolabs), with the number of transformants at least 50 times the possible DNA sequence diversity. Plasmid libraries were purified using GeneJet Maxi Prep kit (ThermoFisher).

Library Selection. Libraries were expressed transiently using Expifectamine (ThermoFisher) in Expi293F cells. To reduce the likelihood of two SERT sequences being acquired by a single cell (*48*), 1ng library DNA was diluted with 1.5 µg pCEP4ΔCMV (*46*) carrier plasmid per milliliter of culture at 2×10^6 cells/ml. Cells were incubated at 37 ºC, 8% CO_2, with continuous shaking at 125 rpm for 2 h, followed by media replacement, and then incubated for a further 22 h. Transport activity was assessed by washing with PBS-BSA (phosphate-buffered saline supplemented with 0.2% bovine serum albumin), and incubating in 0.5 µM APP+ (Aobious) in PBS-BSA for 15 min at room temperature with gentle agitation. Cells were then washed with ice cold PBS-BSA, incubated on ice for 20 m with 1/200 anti-myc-Alexa 647 (clone 9B11; Cell Signaling Technology), washed twice, and then analyzed on an Accuri C6 or BD LSRII cytometer, or sorted on a BD FACS ARIA II. During sorting, the main cell population was gated by forward scatter-A versus -W to exclude doublets and debris. Propidium iodide-negative events were then gated to remove dead cells. The top 50 % of SERT-expressing cells (based on Alexa 647 fluorescence) were collected and frozen at -80 °C. For analysis of transport, the top 10-15 % of APP+ fluorescent cells were collected within the SERT-expression gate. To adequately sample each mutation in the libraries during sorts, at least 100,000 cells were collected, pooling frozen cell pellets from separate

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

days as necessary to create one sample from which cDNA was prepared. Typically, ~8 h of sorting was required for each condition, and transfected cultures were not sorted for longer than 4 hours to maintain viability. Each deep mutational scan experiment consisted of at least two separately sorted and prepared transfected cell samples.

Library Sequencing and Analysis. Total RNA was extracted from the sorted cells (GeneJet kit from ThermoFisher), followed by first strand cDNA synthesis primed with gene-specific oligonucleotides using high-fidelity Accuscript (Agilent). DNA fragments were amplified in two rounds of PCR as previously described (*46*, *86*), adding Illumina adapter sequences and barcodes. Samples were sequenced on an Illumina MiSeq or HiSeq2500 using 2×250 nt paired end read protocol. Data were processed using Enrich (*87*), and commands for analysis are included in the GEO submission. Enrichment ratios for each variant were calculated by dividing the transcript frequency from the selected population by the plasmid DNA frequency in the naive plasmid library. Sequence variants with < 10 reads (equivalent to a frequency of $\sim 3.5 \times 10^{-6}$) in the naïve plasmid library were considered missing. Log_2 enrichment scores are displayed as heat maps using gnuplot (http://www.gnuplot.info/), or displayed on the crystal structure of (s)-citalopram-bound SERT (PDB 5I73) using PyMOL (Shrodinger, LLC).

Variant Selection, Targeted Mutagenesis, and Analysis. Variants for validation by targeted mutagenesis were selected from the deep mutational scan with a $\log_2$ enrichment ratio of >3 for APP+ import and <1 for expression, or were reported human alleles. Primers containing the specific codon change were used for site-directed mutagenesis using overlap extension PCR. SERT variants (500 ng plasmid DNA per ml culture) were transiently transfected into Expi293F cells (2×10^6 cells/ml) using Expifectamine (ThermoFisher). Transfection enhancers were added 16-18 h post-transfection, and cells were harvested for analysis 36 h post-transfection. APP+

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

transport activity was assessed as described for library preparation. Samples were analyzed on an Accuri C6 cytometer. Total and surface protein expression were measured using the BD Fixation/Permeabilization Solution Kit (BD Biosciences) in conjunction with anti-myc-Alexa 647 staining (clone 9B11; Cell Signaling Technology). For testing neurotransmitter selectivity, assays were performed on Expi293F cells incubated with 0.5 μM APP+ and varying concentrations of 5HT, dopamine, or norepinephrine in Ringer's Solution (140 mM NaCl, 5 mM KCl, 1.2 mM $MgCl_2$, 5 mM Glucose, 2 mM $CaCl_2$, 10 mM HEPES, pH 7.5) supplemented with 0.2% bovine serum albumin. Ion dependency was measured by replacing NaCl in Ringer's-BSA solution with Choline-Cl (140 mM) and adjusting the pH to 7.5 with choline-hydroxide solution, or by replacing chloride salts with salts of gluconate, and adjusting pH with KOH.

The mean fluorescence (F_{APP+}) of cells incubated with APP+ was measured by flow cytometry. Background autofluorescence (F_0) was subtracted using SERT-transfected cells without APP+ to calculate $\Delta F = F_{APP+} - F_0$. To assess transport activity (T), ΔF of mock transfected control cells incubated with APP+ was subtracted from ΔF of SERT-expressing cells: $T = \Delta F_{SERT} - \Delta F_{negative}$. For quantifying surface expression, autofluorescence of untransfected cells incubated with Alexa 647-conjugated anti-myc antibody was subtracted from the mean fluorescence units of the SERT-transfected population. All replicates are independent samples.

All Atom MD Simulations. Simulations were initiated from an OF Na^+-bound SERT structure, with Na^+ bound in the Na1 and Na2 sites, obtained from our previous study (*44*). The protein was embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer using the Membrane Builder plugin in CHARMM-GUI (*88*). The terminal chains were capped with acetyl and methyl amide groups. 1 APP+ molecule was added to the simulation box. 150 mM NaCl was added to neutralize the system and mimic physiological conditions. Glu508 was modelled in its

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

protonated form. The disulfide bond between Cys200 and Cys209 of EL2 was present. The system was solvated with TIP3P water molecules (*89*) in a box volume of 75 x 75 x 112 Å^3. Simulations were implemented with the Amber18 package (*90*) using Amber ff14SB(*91*) and GAFF force fields(*92*). The MD system was minimized for 20,000 steps using the conjugate gradient method, heated from 0 to 300 K at NVT, and equilibrated for 40 ns under NPT conditions. All production runs were conducted at constant NPT conditions (300K, 1 atm), periodic boundary conditions, and hydrogen mass repartitioning (*93*). An integration timestep of 4 femtosecond (fs) was used. Temperature (300 K) was maintained using Langevin dynamics with a 1.0 picosecond^{-1} damping coefficient. Pressure (1 atm) was maintained using Monte Carlo barostat with an update interval every 100 simulation steps. The nonbonded distance cutoff was set at 10 Å. Electrostatic interactions were treated with the Particle Mesh Ewald method (*94*), and hydrogen bonds were constrained using SHAKE algorithm (*95*). Simulation snapshots were saved for every 100 ps during production simulations.

As conducted in our previous study, we implemented a Markov state model (MSM)-based adaptive sampling approach (*96*, *97*) to thoroughly sample the conformational landscape of SERT during the APP+ import process. This approach involved an iterative sampling procedure in which multiple simulations were conducted in parallel. Data were then clustered based on a designated metric of the z-position of the symported Na^+, Cl^-, and APP+ using a K-means algorithm. States from the least populated clusters were chosen for the next iteration of simulations. A total of ~370 µs of simulation time was collected.

MSM Construction. Markov state modeling is a statistical approach for discretizing the simulation data into kinetically relevant clusters and calculating the transition probability between each cluster (*58*). MSM were constructed with the pyEMMA 2.5.6 Python package (*98*). To

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

compare the APP+ import process with Na^+ and 5HT-SERT simulations, we constructed the MSM using the same 16 interhelical distances surrounding the channel pore radius and z-components of substrates (APP+, Cl^-, and the symported Na^+) as featurization metrics (Figure S6A). The number of clusters was optimized based on the VAMP scoring function implemented in pyEMMA. 4 time independent components (tICs) and 600 clusters were chosen to construct the MSM (Figure S6B). The lag time of 12 ns was determined to be a Markovian lag time based on implied timescale plots (Figure S6C). To validate the model, we perform a Chapman-Kolmogorov test (*99*) on a 4 macrostates MSM (Figure S7).

Trajectory Processing and Analysis. The CPPTRAJ module (*100*) implemented in AmberTools and MDTraj (*101*) was used for post processing of trajectory data. All simulation data were projected on the coordinate landscape defined by the extracellular gating distance as the closest heavy atom between Arg104 and Glu493 and the intracellular distance as the closest heavy atom between the groups of residues Val86-Ser91, Tyr350 and Val199-Trp282, Glu444. To estimate the error of the free energy landscape, a bootstrapping method was implemented in which 500 independent MSM were constructed using a random selection of 80% of the trajectory dataset (Figure S8). Trajectories and MSM states were visualized with Visual Molecular Dynamics (VMD) (*102*) and PyMOL (Shrodinger, LLC). In-house scripts and the matplotlib python package were used to generate plots.

Clustering the mutational landscape using UMAP. Clustering of the mutational landscape was implemented using the UMAP python package (https://github.com/lmcinnes/) using the following hyperparameters: n_neighbors = 10, min_dist = 0.2, n_components = 2, metric = euclidian. The mutational data were then projected on the two UMAP components and colored based on different biochemical and biophysical properties. Calculation of solvent accessible surface area and root

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

mean square fluctuation was performed on 200 randomly sampled structures from each MSM state. EVmutation scores was calculated using the EVcouplings package (https://github.com/debbiemarkslab/EVcouplings).

Estimating Evolutionary Conservation. Natural sequence conservation was mapped to the SERT structure (PDB 5I73) using the ConSurf server (*103*, *104*). Homologs were searched using CS-BLAST from UNIREF-90 with an E-value cutoff of 0.0001 and 3 iterations. Maximum and minimum identity were set at 95 % and 35 %, respectively. 500 NSS transporter sequences were aligned with MAFFT and conservation scores calculated using Bayesian algorithm of substitution.

Data and Software Availability. Raw deep sequencing data and calculated enrichment ratios are deposited with NCBI's Gene Expression Omnibus under series accession number GSE109499. Command lines for analysis with Enrich are also included. Representative structures from the Markov state model and MD simulations are available on Zenodo at https://doi.org/10.5281/zenodo.4695032.

SUPPORTING INFORMATION

Online Supporting Information includes:

PDF of Supplementary Figures S1 to S8

PyMOL file of SERT colored by conservation score for APP+ transport with overlaid APP+ permeation pathway

Movie S1 – morph of SERT conformational transitions colored by APP+ transport conservation score

Excel file of enrichment ratios from the SERT deep mutational scans

AUTHOR INFORMATION

Corresponding Authors

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

* (E.P.) Email: procko@illinois.edu

* (D.S.) Email: diwakar@illinois.edu

Author Contributions

H.J.Y. and S.K.S. did experimental characterization and deep mutagenesis. M.C.C. and B.S. did MD simulations. E.P. and D.S. supervised and directed research.

Notes

The authors declare they have no conflict of interest.

ACKNOWLEDGEMENTS

Within the UIUC Roy J. Carver Biotechnology Center, Barbara Pilas and Angela Kouris assisted with flow cytometry, and Alvaro Hernandez and Chris Wright assisted with deep sequencing. This work was funded by R21 MH113155 from NIMH to E.P., and D.S. is supported by NSF Early Career Award by NSF MCB 18-45606. The research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois, and as of December, 2019, the National Geospatial-Intelligence Agency. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications.

REFERENCES

1. R. D. Blakely, H. E. Berson, R. T. Fremeau, M. G. Caron, M. M. Peek, H. K. Prince, C. C. Bradley, Cloning and expression of a functional serotonin transporter from rat brain. *Nature*. **354**, 66–70 (1991).

2. B. J. Hoffman, E. Mezey, M. J. Brownstein, Cloning of a serotonin transporter affected by antidepressants. *Science*. **254**, 579–580 (1991).

3. D. L. Murphy, G. R. Uhl, A. Holmes, R. Ren-Patterson, F. S. Hall, I. Sora, S. Detera-Wadleigh, K.-P. Lesch, Experimental gene interaction studies with SERT mutant mice as

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

models for human polygenic and epistatic traits and disorders. *Genes, Brain and Behavior*. **2**, 350–364 (2003).

4. F. Kilic, D. L. Murphy, G. Rudnick, A Human Serotonin Transporter Mutation Causes Constitutive Activation of Transport Activity. *Mol Pharmacol*. **64**, 440–446 (2003).

5. N. Ozaki, D. Goldman, W. H. Kaye, K. Plotnicov, B. D. Greenberg, J. Lappalainen, G. Rudnick, D. L. Murphy, Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. *Molecular Psychiatry*. **8**, 933–936 (2003).

6. H. C. Prasad, J. A. Steiner, J. S. Sutcliffe, R. D. Blakely, Enhanced activity of human serotonin transporter variants associated with autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **364**, 163–173 (2009).

7. J. S. Sutcliffe, R. J. Delahanty, H. C. Prasad, J. L. McCauley, Q. Han, L. Jiang, C. Li, S. E. Folstein, R. D. Blakely, Allelic Heterogeneity at the Serotonin Transporter Locus (SLC6A4) Confers Susceptibility to Autism and Rigid-Compulsive Behaviors. *The American Journal of Human Genetics*. **77**, 265–279 (2005).

8. M. K. Hahn, R. D. Blakely, Monoamine transporter gene structure and polymorphisms in relation to psychiatric and other complex disorders. *The Pharmacogenomics Journal*. **2**, 217–235 (2002).

9. T. B. Usdin, E. Mezey, C. Chen, M. J. Brownstein, B. J. Hoffman, Cloning of the cocaine-sensitive bovine dopamine transporter. *PNAS*. **88**, 11168–11171 (1991).

10. B. Giros, S. el Mestikawy, N. Godinot, K. Zheng, H. Han, T. Yang-Feng, M. G. Caron, Cloning, pharmacological characterization, and chromosome assignment of the human dopamine transporter. *Mol Pharmacol*. **42**, 383–390 (1992).

11. T. Pacholczyk, R. D. Blakely, S. G. Amara, Expression cloning of a cocaine-and antidepressant-sensitive human noradrenaline transporter. *Nature*. **350**, 350–354 (1991).

12. G. Rudnick, W. Sandtner, Serotonin transport in the 21st century. *J Gen Physiol*. **151**, 1248–1264 (2019).

13. G. Rudnick, P. J. Nelson, Platelet 5-hydroxytryptamine transport, an electroneutral mechanism coupled to potassium. *Biochemistry*. **17**, 4739–4742 (1978).

14. P. J. Nelson, G. Rudnick, The role of chloride ion in platelet serotonin transport. *J. Biol. Chem*. **257**, 6151–6155 (1982).

15. K. Schicker, Z. Uzelac, J. Gesmonde, S. Bulling, T. Stockner, M. Freissmuth, S. Boehm, G. Rudnick, H. H. Sitte, W. Sandtner, Unifying Concept of Serotonin Transporter-associated Currents. *J. Biol. Chem*. **287**, 438–445 (2012).

16. P. Mitchell, A General Theory of Membrane Transport From Studies of Bacteria. *Nature*. **180**, 134–136 (1957).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

17. O. Jardetzky, Simple Allosteric Model for Membrane Pumps. *Nature*. **211**, 969 (1966).

18. H. Krishnamurthy, E. Gouaux, X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature*. **481**, 469–474 (2012).

19. J. A. Coleman, D. Yang, Z. Zhao, P.-C. Wen, C. Yoshioka, E. Tajkhorshid, E. Gouaux, Serotonin transporter–ibogaine complexes illuminate mechanisms of inhibition and transport. *Nature*. **569**, 141–145 (2019).

20. L. D. Jayanthi, S. Ramamoorthy, Regulation of monoamine transporters: Influence of psychostimulants and therapeutic antidepressants. *AAPS J*. **7**, E728–E738 (2005).

21. F. Lin, H. A. Lester, S. Mager, Single-channel currents produced by the serotonin transporter and analysis of a mutation affecting ion permeation. *Biophysical Journal*. **71**, 3126–3135 (1996).

22. S. Mager, C. Min, D. J. Henry, C. Chavkintt, B. J. Hoffman, N. Davidson, H. A. Lester, Conducting states of a mammalian serotonin transporter. *Neuron*. **12**, 845–859 (1994).

23. M. W. Quick, Regulating the Conducting States of a Mammalian Serotonin Transporter. *Neuron*. **40**, 537–549 (2003).

24. A. S. Kristensen, J. Andersen, T. N. Jørgensen, L. Sørensen, J. Eriksen, C. J. Loland, K. Strømgaard, U. Gether, SLC6 Neurotransmitter Transporters: Structure, Function, and Regulation. *Pharmacol Rev*. **63**, 585–640 (2011).

25. A. Yamashita, S. K. Singh, T. Kawate, Y. Jin, E. Gouaux, Crystal structure of a bacterial homologue of Na + /Cl - -dependent neurotransmitter transporters. *Nature*. **437**, 215 (2005).

26. J. A. Coleman, E. M. Green, E. Gouaux, X-ray structures and mechanism of the human serotonin transporter. *Nature*. **532**, 334–339 (2016).

27. J. A. Coleman, E. Gouaux, Structural basis for recognition of diverse antidepressants by the human serotonin transporter. *Nature Structural & Molecular Biology*. **25**, 170 (2018).

28. A. Penmatsa, K. H. Wang, E. Gouaux, X-ray structures of Drosophila dopamine transporter in complex with nisoxetine and reboxetine. *Nature Structural & Molecular Biology*. **22**, 506–508 (2015).

29. A. Penmatsa, K. H. Wang, E. Gouaux, X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature*. **503**, 85–90 (2013).

30. K. H. Wang, A. Penmatsa, E. Gouaux, Neurotransmitter and psychostimulant recognition by the dopamine transporter. *Nature*. **521**, 322–327 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

31. H. Wang, A. Goehring, K. H. Wang, A. Penmatsa, R. Ressler, E. Gouaux, Structural basis for action by diverse antidepressants on biogenic amine transporters. *Nature*. **503**, 141–145 (2013).

32. L. Sørensen, J. Andersen, M. Thomsen, S. M. R. Hansen, X. Zhao, A. Sandelin, K. Strømgaard, A. S. Kristensen, Interaction of antidepressants with the serotonin and norepinephrine transporters: mutational studies of the S1 substrate binding pocket. *J. Biol. Chem*. **287**, 43694–43707 (2012).

33. Z. Zhou, J. Zhen, N. K. Karpowich, R. M. Goetz, C. J. Law, M. E. A. Reith, D.-N. Wang, LeuT-Desipramine Structure Reveals How Antidepressants Block Neurotransmitter Reuptake. *Science*. **317**, 1390–1393 (2007).

34. S. K. Singh, A. Yamashita, E. Gouaux, Antidepressant binding site in a bacterial homologue of neurotransmitter transporters. *Nature*. **448**, 952–956 (2007).

35. H. A. Neubauer, C. G. Hansen, O. Wiborg, Dissection of an Allosteric Mechanism on the Serotonin Transporter: A Cross-Species Study. *Mol Pharmacol*. **69**, 1242–1250 (2006).

36. L. Malinauskaite, M. Quick, L. Reinhard, J. A. Lyons, H. Yano, J. A. Javitch, P. Nissen, A mechanism for intracellular release of Na+ by neurotransmitter/sodium symporters. *Nat. Struct. Mol. Biol*. **21**, 1006–1012 (2014).

37. Y. Zhao, D. Terry, L. Shi, H. Weinstein, S. C. Blanchard, J. A. Javitch, Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature*. **465**, 188–193 (2010).

38. Y. Zhao, D. S. Terry, L. Shi, M. Quick, H. Weinstein, S. C. Blanchard, J. A. Javitch, Substrate-modulated gating dynamics in a Na^+-coupled neurotransmitter transporter homologue. *Nature*. **474**, 109–113 (2011).

39. L. R. Forrest, Y.-W. Zhang, M. T. Jacobs, J. Gesmonde, L. Xie, B. H. Honig, G. Rudnick, Mechanism for alternating access in neurotransmitter transporters. *PNAS*. **105**, 10338–10343 (2008).

40. A. M. Razavi, G. Khelashvili, H. Weinstein, A Markov State-based Quantitative Kinetic Model of Sodium Release from the Dopamine Transporter. *Scientific Reports*. **7**, 40076 (2017).

41. S. A. Shaikh, E. Tajkhorshid, Modeling and Dynamics of the Inward-Facing State of a Na+/Cl− Dependent Neurotransmitter Transporter Homologue. *PLOS Computational Biology*. **6**, e1000905 (2010).

42. S. Adhikary, D. J. Deredge, A. Nagarajan, L. R. Forrest, P. L. Wintrode, S. K. Singh, Conformational dynamics of a neurotransmitter:sodium symporter in a lipid bilayer. *PNAS*. **114**, E1786–E1795 (2017).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

43. I. R. Möller, M. Slivacka, A. K. Nielsen, S. G. F. Rasmussen, U. Gether, C. J. Loland, K. D. Rand, Conformational dynamics of the human serotonin transporter during substrate and drug binding. *Nature Communications*. **10**, 1687 (2019).

44. M. Chan, B. Selvam, H. Young, E. Procko, D. Shukla, The Substrate Import Mechanism of the Human Serotonin Transporter (2020), doi:10.26434/chemrxiv.9922301.v2.

45. D. M. Fowler, S. Fields, Deep mutational scanning: a new style of protein science. *Nature Methods*. **11**, 801–807 (2014).

46. J. Park, B. Selvam, K. Sanematsu, N. Shigemura, D. Shukla, E. Procko, Structural architecture of a dimeric class C GPCR based on co-trafficking of sweet taste receptor subunits. *J. Biol. Chem.* **294**, 4759–4774 (2019).

47. J. D. Heredia, J. Park, H. Choi, K. S. Gill, E. Procko, Conformational Engineering of HIV-1 Env Based on Mutational Tolerance in the CD4 and PG16 Bound States. *Journal of Virology*. **93** (2019), doi:10.1128/JVI.00219-19.

48. J. D. Heredia, J. Park, R. J. Brubaker, S. K. Szymanski, K. S. Gill, E. Procko, Mapping Interaction Sites on Human Chemokine Receptors by Deep Mutational Scanning. *The Journal of Immunology*. **200**, 3825–3839 (2018).

49. A. C. McShan, C. A. Devlin, S. A. Overall, J. Park, J. S. Toor, D. Moschidi, D. Flores-Solis, H. Choi, S. Tripathi, E. Procko, N. G. Sgourakis, Molecular determinants of chaperone interactions on MHC-I for folding and antigen repertoire selection. *PNAS*. **116**, 25602–25613 (2019).

50. N. J. Rollins, K. P. Brock, F. J. Poelwijk, M. A. Stiffler, N. P. Gauthier, C. Sander, D. S. Marks, Inferring protein 3D structure from deep mutation scans. *Nat Genet*. **51**, 1170–1176 (2019).

51. J. M. Schmiedel, B. Lehner, Determining protein structures using deep mutagenesis. *Nature Genetics*. **51**, 1177–1186 (2019).

52. B. Felts, A. B. Pramod, W. Sandtner, N. Burbach, S. Bulling, H. H. Sitte, L. K. Henry, The Two Na+ Sites in the Human Serotonin Transporter Play Distinct Roles in the Ion Coupling and Electrogenicity of Transport. *J Biol Chem*. **289**, 1825–1840 (2014).

53. S. Ramamoorthy, A. L. Bauman, K. R. Moore, H. Han, T. Yang-Feng, A. S. Chang, V. Ganapathy, R. D. Blakely, Antidepressant- and cocaine-sensitive human serotonin transporter: molecular cloning, expression, and chromosomal localization. *Proc Natl Acad Sci U S A*. **90**, 2542–2546 (1993).

54. G. Rudnick, Active transport of 5-hydroxytryptamine by plasma membrane vesicles isolated from human blood platelets. *J. Biol. Chem.* **252**, 2170–2174 (1977).

55. C. Fenollar-Ferrer, T. Stockner, T. C. Schwarz, A. Pal, J. Gotovina, T. Hofmaier, K. Jayaraman, S. Adhikary, O. Kudlacek, A. R. Mehdipour, S. Tavoulari, G. Rudnick, S. K.

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

Singh, R. Konrat, H. H. Sitte, L. R. Forrest, Structure and Regulatory Interactions of the Cytoplasmic Terminal Domains of Serotonin Transporter. *Biochemistry*. **53**, 5444–5460 (2014).

56. H. C. Prasad, C.-B. Zhu, J. L. McCauley, D. J. Samuvel, S. Ramamoorthy, R. C. Shelton, W. A. Hewlett, J. S. Sutcliffe, R. D. Blakely, Human serotonin transporter variants display altered sensitivity to protein kinase G and p38 mitogen-activated protein kinase. *PNAS*. **102**, 11545–11550 (2005).

57. Y.-W. Zhang, J. Gesmonde, S. Ramamoorthy, G. Rudnick, Serotonin Transporter Phosphorylation by cGMP-Dependent Protein Kinase Is Altered by a Mutation Associated with Obsessive–Compulsive Disorder. *J. Neurosci*. **27**, 10878–10886 (2007).

58. B. E. Husic, V. S. Pande, Markov State Models: From an Art to a Science. *J. Am. Chem. Soc*. **140**, 2386–2396 (2018).

59. B. Selvam, S. Mittal, D. Shukla, Free Energy Landscape of the Complete Transport Cycle in a Key Bacterial Transporter. *ACS Cent Sci*. **4**, 1146–1154 (2018).

60. D. Shukla, C. X. Hernández, J. K. Weber, V. S. Pande, Markov State Models Provide Insights into Dynamic Modulation of Protein Function. *Acc. Chem. Res*. **48**, 414–422 (2015).

61. B. Selvam, Y.-C. Yu, L.-Q. Chen, D. Shukla, Molecular Basis of the Glucose Transport Mechanism in Plants. *ACS Cent. Sci*. **5**, 1085–1096 (2019).

62. D. P. Claxton, M. Quick, L. Shi, F. D. de Carvalho, H. Weinstein, J. A. Javitch, H. S. Mchaourab, Ion/substrate-dependent conformational dynamics of a bacterial homolog of neurotransmitter:sodium symporters. *Nature Structural & Molecular Biology*. **17**, 822–829 (2010).

63. M. Lek, K. J. Karczewski, E. V. Minikel, K. E. Samocha, E. Banks, T. Fennell, A. H. O'Donnell-Luria, J. S. Ware, A. J. Hill, B. B. Cummings, T. Tukiainen, D. P. Birnbaum, J. A. Kosmicki, L. E. Duncan, K. Estrada, F. Zhao, J. Zou, E. Pierce-Hoffman, J. Berghout, D. N. Cooper, N. Deflaux, M. DePristo, R. Do, J. Flannick, M. Fromer, L. Gauthier, J. Goldstein, N. Gupta, D. Howrigan, A. Kiezun, M. I. Kurki, A. L. Moonshine, P. Natarajan, L. Orozco, G. M. Peloso, R. Poplin, M. A. Rivas, V. Ruano-Rubio, S. A. Rose, D. M. Ruderfer, K. Shakir, P. D. Stenson, C. Stevens, B. P. Thomas, G. Tiao, M. T. Tusie-Luna, B. Weisburd, H.-H. Won, D. Yu, D. M. Altshuler, D. Ardissino, M. Boehnke, J. Danesh, S. Donnelly, R. Elosua, J. C. Florez, S. B. Gabriel, G. Getz, S. J. Glatt, C. M. Hultman, S. Kathiresan, M. Laakso, S. McCarroll, M. I. McCarthy, D. McGovern, R. McPherson, B. M. Neale, A. Palotie, S. M. Purcell, D. Saleheen, J. M. Scharf, P. Sklar, P. F. Sullivan, J. Tuomilehto, M. T. Tsuang, H. C. Watkins, J. G. Wilson, M. J. Daly, D. G. MacArthur, Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. **536**, 285–291 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

64. S. Sucic, S. Dallinger, B. Zdrazil, R. Weissensteiner, T. N. Jørgensen, M. Holy, O. Kudlacek, S. Seidel, J. H. Cha, U. Gether, A. H. Newman, G. F. Ecker, M. Freissmuth, H. H. Sitte, The N Terminus of Monoamine Transporters Is a Lever Required for the Action of Amphetamines. *J. Biol. Chem.* **285**, 10924–10938 (2010).

65. P. Metzner, C. Schütte, E. Vanden-Eijnden, Transition Path Theory for Markov Jump Processes. *Multiscale Model. Simul.* **7**, 1192–1219 (2009).

66. P. C. Söderhielm, J. Andersen, L. Munro, A. T. Nielsen, A. S. Kristensen, Substrate and Inhibitor–Specific Conformational Changes in the Human Serotonin Transporter Revealed by Voltage-Clamp Fluorometry. *Mol Pharmacol.* **88**, 676–688 (2015).

67. S. Tavoulari, E. Margheritis, A. Nagarajan, D. C. DeWitt, Y.-W. Zhang, E. Rosado, S. Ravera, E. Rhoades, L. R. Forrest, G. Rudnick, Two Na+ Sites Control Conformational Change in a Neurotransmitter Transporter Homolog. *J. Biol. Chem.* **291**, 1456–1471 (2016).

68. J. Andersen, L. Olsen, K. B. Hansen, O. Taboureau, F. S. Jørgensen, A. M. Jørgensen, B. Bang-Andersen, J. Egebjerg, K. Strømgaard, A. S. Kristensen, Mutational mapping and modeling of the binding site for (S)-citalopram in the human serotonin transporter. *J. Biol. Chem.* **285**, 2051–2063 (2010).

69. H. Koldsø, K. Severinsen, T. T. Tran, L. Celik, H. H. Jensen, O. Wiborg, B. Schiøtt, S. Sinning, The two enantiomers of citalopram bind to the human serotonin transporter in reversed orientations. *J. Am. Chem. Soc.* **132**, 1311–1322 (2010).

70. L. McInnes, J. Healy, J. Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426 [cs, stat]* (2020) (available at http://arxiv.org/abs/1802.03426).

71. M. W. Dorrity, L. M. Saunders, C. Queitsch, S. Fields, C. Trapnell, Dimensionality reduction by UMAP to visualize physical and genetic interactions. *Nature Communications.* **11**, 1537 (2020).

72. E. M. Jones, N. B. Lubock, A. Venkatakrishnan, J. Wang, A. M. Tseng, J. M. Paggi, N. R. Latorraca, D. Cancilla, M. Satyadi, J. E. Davis, M. M. Babu, R. O. Dror, S. Kosuri, Structural and functional characterization of G protein–coupled receptors with deep mutational scanning. *eLife.* **9**, e54895 (2020).

73. T. A. Hopf, J. B. Ingraham, F. J. Poelwijk, C. P. I. Schärfe, M. Springer, C. Sander, D. S. Marks, Mutation effects predicted from sequence co-variation. *Nat Biotechnol.* **35**, 128–135 (2017).

74. K. Kazmier, S. Sharma, S. M. Islam, B. Roux, H. S. Mchaourab, Conformational cycle and ion-coupling mechanism of the Na+/hydantoin transporter Mhp1. *PNAS.* **111**, 14752–14757 (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

75. E. Solis, I. Zdravkovic, I. D. Tomlinson, S. Y. Noskov, S. J. Rosenthal, L. J. De Felice, 4-(4-(dimethylamino)phenyl)-1-methylpyridinium (APP+) is a fluorescent substrate for the human serotonin transporter. *J. Biol. Chem.* **287**, 8852–8863 (2012).

76. C. A. Tracewell, F. H. Arnold, Directed enzyme evolution: climbing fitness peaks one amino acid at a time. *Current Opinion in Chemical Biology*. **13**, 3–9 (2009).

77. A. Reinders, Y. Sun, K. L. Karvonen, J. M. Ward, Identification of Amino Acids Important for Substrate Specificity in Sucrose Transporters Using Gene Shuffling. *J. Biol. Chem.* **287**, 30296–30304 (2012).

78. A. Farwick, S. Bruder, V. Schadeweg, M. Oreb, E. Boles, Engineering of yeast hexose transporters to transport d-xylose without inhibition by d-glucose. *PNAS*. **111**, 5159–5164 (2014).

79. J. G. Nijland, H. Y. Shin, R. M. de Jong, P. P. de Waal, P. Klaassen, A. J. Driessen, Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in Saccharomyces cerevisiae. *Biotechnology for Biofuels*. **7**, 168 (2014).

80. E. M. Young, A. Tong, H. Bui, C. Spofford, H. S. Alper, Rewiring yeast sugar transporter preference through modifying a conserved protein motif. *PNAS*. **111**, 131–136 (2014).

81. C. L. Piscitelli, E. Gouaux, Insights into transport mechanism from LeuT engineered to transport tryptophan. *The EMBO Journal*. **31**, 228–235 (2012).

82. A. Anderluh, E. Klotzsch, A. W. A. F. Reismann, M. Brameshuber, O. Kudlacek, A. H. Newman, H. H. Sitte, G. J. Schütz, Single molecule analysis reveals coexistence of stable serotonin transporter monomers and oligomers in the live cell plasma membrane. *J. Biol. Chem.* **289**, 4387–4394 (2014).

83. A. Anderluh, T. Hofmaier, E. Klotzsch, O. Kudlacek, T. Stockner, H. H. Sitte, G. J. Schütz, Direct PIP$_2$ binding mediates stable oligomer formation of the serotonin transporter. *Nature Communications*. **8**, 14089 (2017).

84. X. Periole, T. Zeppelin, B. Schiøtt, Dimer Interface of the Human Serotonin Transporter and Effect of the Membrane Composition. *Scientific Reports*. **8**, 5080 (2018).

85. M. A. Ciccone, M. Timmons, A. Phillips, M. W. Quick, Calcium/Calmodulin-Dependent Kinase II Regulates the Interaction between the Serotonin Transporter and Syntaxin 1A. *Neuropharmacology*. **55**, 763–770 (2008).

86. E. Procko, R. Hedman, K. Hamilton, J. Seetharaman, S. J. Fleishman, M. Su, J. Aramini, G. Kornhaber, J. F. Hunt, L. Tong, G. T. Montelione, D. Baker, Computational Design of a Protein-Based Enzyme Inhibitor. *Journal of Molecular Biology*. **425**, 3563–3575 (2013).

87. D. M. Fowler, C. L. Araya, W. Gerard, S. Fields, Enrich: software for analysis of protein function by enrichment and depletion of variants. *Bioinformatics*. **27**, 3430–3431 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

88. S. Jo, T. Kim, V. G. Iyer, W. Im, CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem*. **29**, 1859–1865 (2008).

89. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

90. R. Salomon-Ferrer, D. A. Case, R. C. Walker, An overview of the Amber biomolecular simulation package. *WIREs Computational Molecular Science*. **3**, 198–210 (2013).

91. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput*. **11**, 3696–3713 (2015).

92. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J Comput Chem*. **25**, 1157–1174 (2004).

93. C. W. Hopkins, S. Le Grand, R. C. Walker, A. E. Roitberg, Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).

94. T. Darden, D. York, L. Pedersen, Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

95. V. Kräutler, W. F. van Gunsteren, P. H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry*. **22**, 501–508 (2001).

96. G. R. Bowman, D. L. Ensign, V. S. Pande, Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. *J. Chem. Theory Comput.* **6**, 787–794 (2010).

97. M. I. Zimmerman, J. R. Porter, X. Sun, R. R. Silva, G. R. Bowman, Choice of Adaptive Sampling Strategy Impacts State Discovery, Transition Probabilities, and the Apparent Mechanism of Conformational Changes. *J. Chem. Theory Comput.* **14**, 5459–5475 (2018).

98. M. K. Scherer, B. Trendelkamp-Schroer, F. Paul, G. Pérez-Hernández, M. Hoffmann, N. Plattner, C. Wehmeyer, J.-H. Prinz, F. Noé, PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J. Chem. Theory Comput.* **11**, 5525–5542 (2015).

99. J.-H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte, F. Noé, Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **134**, 174105 (2011).

100. D. R. Roe, T. E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2021.04.19.440442; this version posted April 19, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-ND 4.0 International license.

101. R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernández, C. R. Schwantes, L.-P. Wang, T. J. Lane, V. S. Pande, MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal*. **109**, 1528–1532 (2015).

102. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *Journal of Molecular Graphics*. **14**, 33–38 (1996).

103. H. Ashkenazy, E. Erez, E. Martz, T. Pupko, N. Ben-Tal, ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*. **38**, W529–W533 (2010).

104. G. Celniker, G. Nimrod, H. Ashkenazy, F. Glaser, E. Martz, I. Mayrose, T. Pupko, N. Ben-Tal, ConSurf: Using Evolutionary Data to Raise Testable Hypotheses about Protein Function. *Israel Journal of Chemistry*. **53**, 199–206 (2013).