Energy landscape of LeuT from molecular simulations

Mert Gur,^{1,2} Elia Zomot,¹ Mary Hongying Cheng,¹ and Ivet Bahar^{1,a)}

¹*Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA*

²*Department of Mechanical Engineering, Istanbul Technical University (ITU), Istanbul 34437, Turkey*

(Received 11 September 2015; accepted 6 November 2015; published online 24 November 2015)

The bacterial sodium-coupled leucine transporter (LeuT) has been broadly used as a structural model for understanding the structure-dynamics-function of mammalian neurotransmitter transporters as well as other solute carriers that share the same fold (LeuT fold), as the first member of the family crystallographically resolved in multiple states: outward-facing open, outward-facing occluded, and inward-facing open. Yet, a complete picture of the energy landscape of (sub)states visited along the LeuT transport cycle has been elusive. In an attempt to visualize the conformational spectrum of LeuT, we performed extensive simulations of LeuT dimer dynamics in the presence of substrate (Ala or Leu) and co-transported Na⁺ ions, in explicit membrane and water. We used both conventional molecular dynamics (MD) simulations (with Anton supercomputing machine) and a recently introduced method, *collective* MD, that takes advantage of collective modes of motions predicted by the anisotropic network model. Free energy landscapes constructed based on ~40 μ s trajectories reveal multiple substates occluded to the extracellular (EC) and/or intracellular (IC) media, varying in the levels of exposure of LeuT to EC or IC vestibules. The IC-facing transmembrane (TM) helical segment TM1a shows an opening, albeit to a smaller extent and in a slightly different direction than that observed in the inward-facing open crystal structure. The study provides insights into the spectrum of conformational substates and paths accessible to LeuT and highlights the differences between Ala- and Leu-bound substates. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4936133>]

I. INTRODUCTION

Neurotransmitters are small signaling molecules that facilitate the communication between neurons. They transmit signals across the specialized junctions, called synapses, connecting the nerve terminals of presynaptic neurons and the dendrites of postsynaptic neurons. Upon the arrival of a signal, neurotransmitters are released in a calcium-dependent manner from the presynaptic neurons into the synaptic gap. Released neurotransmitters bind and activate ligand-gated ion channels and G-protein-coupled receptors on the postsynaptic neurons. These activated proteins, in turn, generate and modulate excitatory (or inhibitory) signals and flows in the postsynaptic neurons.¹ The signal is thus chemically transmitted from the presynaptic neurons to the postsynaptic ones across the synapse. Due to the critical function of neurotransmitters in the central and peripheral nervous systems, it is of utmost importance to control their spatiotemporal level in the synapse. Neurotransmitter transporters control these levels by reuptake of excess neurotransmitters from the synapse into presynaptic cells or glial cells.^{2,3} These secondary transporters typically use the flow of Na⁺ (with or without Cl⁻ ions) down its electrochemical gradient across the cell membrane to drive the (co)transport of neurotransmitters (or substrate) against their concentration gradient,⁴ hence their family name neurotransmitter:sodium symporter (NSS). The NSS

family includes members that transport dopamine, serotonin, noradrenaline, GABA, and glycine. NSS dysfunctions are related to numerous neurological disorders such as epilepsy,⁵ depression,^{6,7} anxiety,^{8,9} attention-deficit hyperactivity disorder,¹⁰ orthostatic intolerance,¹¹ and Parkinson's disease.⁹ Thus, NSS members are targets for a wide range of drugs against various neurological disorders. Despite side effects,¹² drugs that inhibit neurotransmitter uptake have been broadly adopted for treating depression.^{13,14} Imipramine (Tofranil), which inhibits noradrenaline (norepinephrine) uptake,¹⁵ or fluoxetine (Prozac), which inhibits serotonin uptake,¹⁶ is an example of such clinical drugs. NSS members are also targeted by illicit drugs such as cocaine and amphetamine.^{17,18}

The sodium:leucine transporter (LeuT) is the first crystallographically resolved member of the NSS family, making LeuT a model for studying homologous mammalian members of the family. The originally resolved LeuT structure¹⁹ was in *outward-facing* (OF) *occluded* (or *closed*) (OFc) state. Over the past years, LeuT has continued to serve as a model, especially for understanding the structure-dynamics-function relationships in the NSS family, due to the large number of crystal structures resolved in its various states: *outward-facing open* (OFO), OFc, and *inward-facing open* (IFO). In the OF state (open or occluded), LeuT structures were crystallized as homodimers.

The so-called *LeuT fold* of the LeuT monomers is shared with other Na⁺-coupled transporters including dopamine transporter (DAT), Mhp1, vSGLT, BetP, CaiT, and AdiC1. The fold is composed of 12 transmembrane (TM) helices

^{a)}Author to whom correspondence should be addressed. Electronic mail: bahar@pitt.edu. Tel.: 412 6483332. Fax: 412 6483163. URL: <http://www.ccbb.pitt.edu/Faculty/bahar/>.

organized in two 5-helix inverted repeats: TM helices 1-5 and 6-10 can be positioned on a pseudo twofold axis with respect to each other.¹⁹ The α -helical geometry of the first TM helix in each of these repeated substructures, TM1 and TM6, is disrupted around halfway across the membrane exposing carbonyl/amide groups that bind the substrate and two Na⁺ ions. The canonical transition between the OF_o and IF_o states involves major reorientations in TM1 and TM6. These, accompanied by other cooperative rearrangements, expose the substrate-binding pocket either to the extracellular (EC) medium for substrate (re)uptake or to the intracellular (IC) environment for substrate release. It is generally accepted that NSSs function via alternating access mechanism,^{20–22} while the detailed functional mechanism at full atomic scale is still a topic of current research.

The OF_o LeuT structure has been resolved in both a substrate-free form as a double mutant Y108F/K288A (PDB ID: 3TT1²³) and a competitive inhibitor-bound wild type (WT) form (PDB ID: 3F3A²⁴). Structures in the OF_c state have been resolved in substrate- and inhibitor-bound forms with the E290S mutation (PDB ID: 3GJC²⁵), in substrate- and noncompetitive inhibitor-bound WT form (PDB ID: 2Q6H²⁶), and in substrate-bound WT forms (PDB ID: 2A65¹⁹ and PDB ID: 3F3E²⁴). The OF_c form is distinguished from the OF_o mainly by $\sim 10^\circ$ reorientations in the EC halves of TM1, 2, and 6 towards the center, and by a $\sim 90^\circ$ rotation in the χ_1 dihedral of F253, which together with Y108 form the EC thin gate.²¹ These changes obstruct the access of the substrate-binding site to the EC environment, hence the name occluded.^{23,24} A detailed description of LeuT sodium and substrate binding and unbinding properties on the OF state can be found in our recent 20 μ s molecular dynamics (MD) simulations of the OF_c and OF_o states in their apo and holo forms.²⁷

In contrast to the OF states, only a single crystal structure, apo IF_o, exists for the IF state of LeuT. In order to stabilize and be able to crystallize the IF structure, three mutations were introduced into the K288A-LeuT background: ((i) and (ii)) the T354V and S355A double mutation, which abolishes sodium binding at the Na2 site, and (iii) the Y268A mutation, which weakens the closure of the IC gate.²³ The major differences observed therein in the IC state with respect to the OF state were the outward rotations of the IC halves of TM1 (TM1a) and 6 (TM6b) by $\sim 45^\circ$ and $\sim 17^\circ$, respectively. In addition, the EC portions of TM1 and 6 were tilted inward by $\sim 24^\circ$ and $\sim 21^\circ$, respectively, in the IF_o state, compared to the OF_c state (PDB ID: 3F3E²⁴), further closing the EC vestibule. These rearrangements clearly block the accessibility of substrate and sodium ions to the EC vestibule while increasing their exposure to the IC vestibule. However, whether, or how accurately, these mutant structures are representative of the IF conformation of the WT transporter under physiological conditions, is yet to be clarified.

Another major question of interest is whether the OF_o, OF_c, and IF_o structures represent all stable structures, i.e., free energy minima, along the LeuT transport cycle. Experimental and/or computational information on additional (sub)states and intermediates are scarce in the literature. Recently, using a combination of accelerated, targeted, and conventional MD, we have identified new *apo* and *holo* intermediates, where the

substrate/sodium binding sites are inaccessible from either the EC or IC media.²⁸ However, it is currently not established whether other stable states exist along the canonical LeuT transport cycle or how easily these states would be accessible in the absence of accelerated MD runs that facilitate larger scale excursions in the conformational space. To the best of our knowledge, the energy landscape of LeuT, spanning the vicinity of the OF and IF states, remains uncharted territory. Its potential dependence on the type of substrate, Ala or Leu, also remains to be elucidated.

In this study, we address these issues upon performing $\sim 40 \mu$ s of unbiased MD simulations for LeuT dimer in the presence of explicit water and membrane, shortly referred to as conventional MD (and designated as cMD). Whether the biological assembly of LeuT is dimeric is an open issue, although the crystal structures for the OF state have been consistently resolved in the dimeric form. Because the simulations are performed for the dimer, the simulations actually provide a total of $\sim 80 \mu$ s trajectory for analyzing monomer dynamics, in addition to giving insights into inter-monomer couplings. All simulations were performed under physiological transition conditions for the apo, and the Ala/2Na⁺- and Leu/2Na⁺- bound forms, and 59% of them were initiated from the known crystal structures. The remaining were initiated using conformers generated by *collective* molecular dynamics (coMD)²⁹—a recently developed methodology that takes advantage of the collective modes of motions intrinsically encoded by the fold, while evaluating the interactions and energetics via a full-atomic MD simulation protocol. The collective modes of motion are obtained using the Anisotropic Network Model (ANM).³⁰

Our simulations sampled the conformational space near the structurally resolved states as well as the paths connecting these states, thus providing an extensive description of the energy landscape for the OF $\rightleftharpoons$ IF transition of LeuT. Strikingly, we identified at least two stable minima in the region between the OF_c and IF_o states, where the primary substrate-binding cavity was occluded to both sides of the membrane, consistent with our earlier MD simulations.²⁸ Simulations yielded a first map of the energy landscape being sampled by the LeuT during its transport function and indicated a dependency on the transported substrate (Leu or Ala). Our simulations also show that TM1a opens up in the IF_o state; but the direction of opening differs the size remains smaller compared to that observed in the crystal structure.

II. MATERIALS AND METHODS

A. System preparation

The PDB files 3F3E, 3TT1, and 3TT3 described above are used as initial conformers for the respective OF_c, OF_o, and IF_o states. The former (Leu-bound) is highly similar to the OF_c structure crystallized in the presence of other substrates such as alanine and methionine (e.g., backbone root-mean-square deviation (RMSD) ~ 0.3 Å from PDB ID: 2A65). The OF_o structure (3TT1) is highly similar (backbone RMSD of ~ 0.5 Å) to the OF_o structure resolved in the presence of a competitive inhibitor (tryptophan) (PDB ID: 3F3A).

WT forms of the OFo and IFo structures were generated *in silico*, by replacing the mutated residues by their wild-type counterparts. Each LeuT structure was then embedded in a 1-palmitoyl-2-oleoyl-phosphatidylethanolamine (POPE) membrane bilayer of size $155 \times 110 \text{ \AA}^2$ and aqueous solution with 150 mM NaCl. Water (TIP3P) molecules and POPE head groups were modeled explicitly, whereas united-atom model was used for the POPE acyl chains.³¹ Each system contained $\sim 118\,000$ atoms in total, including $\sim 16\,000$ for protein, $\sim 20\,000$ for the membrane, $\sim 82\,000$ for water, and ~ 164 ions (77 Na^+ and 87 Cl^-). Every equilibrated system had a padding of at least $19 \text{ \AA}$ of lipid molecules (along the x/y axes) or $18 \text{ \AA}$ of water molecules (along the z axis), in each direction such that the minimal distance between the protein and each image was twice these values. Titratable residues were left in the dominant protonation state at pH 7.0. Table I presents the complete list of trajectories analyzed in the present study.

B. Simulation protocol

All simulations were performed using the CHARMM36 force field for lipids and proteins (with the CMAP correction).^{32,33} Structures were equilibrated for 30 ns with harmonic constraints on the protein backbone and on substrate/ Na^+ atoms using a force constant of $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for the first 10 ns and $4 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ for the following by 20 ns. After this initial equilibration step, residues were substituted (from mutant to wild type), the substrate and Na^+ were inserted or removed depending on the initial conditions

(see Table I), the system was neutralized, and a 2nd energy minimization of 20 ns with softer force constant ($2 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ at the same atoms) was performed to allow the systems to adapt to the introduced changes before initiating the productive runs. Each system was then run constraint-free for 20 ns using NAMD-2.8³⁴ before transferring the data to the 512-node Anton machine³⁵ for extensive simulations. A time step of 2 fs was used for calculating short-range electrostatic and van der Waals interactions (cutoff distance of $12 \text{ \AA}$), and 4 fs for electrostatics using the Particle Mesh Ewald method. All systems were simulated in an NPT ensemble at 1 bar and 310 K, using the Berendsen coupling scheme. Trajectory visualizations and analyses were performed using VMD-1.9.1-2.³⁶ Targeted MD (tMD)³⁷ runs, unless indicated otherwise, were performed with a spring constant of $500 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$.

C. Principal component analysis (PCA) of crystal structures

Comparison of LeuT crystal structures showed that TM3 and TM8 underwent minimal movements (RMSD of $0.95 \text{ \AA}$) during the transition between the OFo and IFo states. Hence, the structures as well as snapshots from simulations were structurally aligned upon optimal superposition the C^α -atoms on these two helices. PCA was performed using the three LeuT crystal structures. To this aim we constructed the $3N \times 3N$ covariance matrix

$$\mathbf{C} = \langle (\mathbf{R} - \langle \mathbf{R} \rangle) (\mathbf{R} - \langle \mathbf{R} \rangle)^T \rangle, \quad (1)$$

TABLE I. Initial simulation conditions and durations of the runs performed.

Run ID	Initial state (PDB structure)	Form	LeuT sequence	Initially bound substrate/ion	Simulation duration (μs)
1	OF open (3TT1)	Dimer	WT	2Na^+	1.08
2a-b 3	OF occluded (3F3E)	Dimer	K288A WT	None	(a) 1.05, (b) 1.05 2.22
4 5a-c	OF open (3TT1)	Dimer	WT	Ala/ 2Na^+ Leu/ 2Na^+	0.55 (a) 1.94, (b) 1.51, (c) 1.08
6 7a-c	OF occluded (3F3E)	Dimer	WT	Ala/ 2Na^+ Leu/ 2Na^+	2.63 (a) 3.37, (b) 3.03, (c) 1.59
8a-b	IF open (3TT3)	Dimer	WT	None	(a) 1.07, (b) 1.07
9	co-path conformer 29	Dimer	WT	None	2.21
10	co-path conformer 30	Dimer	WT	None	2.55
11	co-path conformer 29	Dimer	WT	Ala/ 2Na^+	2.37
12	co-path conformer 30	Dimer	WT	Ala/ 2Na^+	2.88
13	co-path conformer 27	Dimer	WT	Leu/ 2Na^+	2.76
14	co-path conformer 28	Dimer	WT	Leu/ 2Na^+	4.68
15	OF open (3TT1)	Monomer	K288A	Ala/ 2Na^+	0.145
16	OF occluded (2A65)	Monomer	K288A	Ala/ 2Na^+	0.08
17	IF open (3TT3)	Monomer	K288A	None	0.18

where $\mathbf{R}$ is the $3N$ -dimensional configuration vector for LeuT in a given crystal structure. $\mathbf{R}$ is composed of the resolved coordinates for the C^α -atoms of residues R11-A504 ($N = 493$); $\langle \mathbf{R} \rangle$ is the average over the three $\mathbf{R}$ vectors. Principal components (PCs) are obtained by eigenvalue decomposition of $\mathbf{C}$ as

$$\mathbf{C} = \sum_{i=1}^{3N} \sigma_i \mathbf{P}_i \mathbf{P}_i^T, \quad (2)$$

where $\mathbf{P}_i$ is the i th principal direction/axis of structural variation from the average and σ_i is the corresponding variance (accounting for the size of variation along that axis). The first PC, $\mathbf{P}_1$ (or PC1), and its variance, σ_1 , define the shape and size of the most dominant structural change. The PCA permitted us to identify three principal axes, two of which were then used for defining the axes of the energy landscape onto which MD snapshots were projected.

D. Anisotropic network model

The ANM³⁰ uses a potential of the form

$$V_{\text{ANM}} = -\frac{\gamma}{2} \left[\sum_{i=1}^{N-1} \sum_{j=i+1}^N (R_{ij} - R_{ij}^0)^2 \Gamma_{ij} \right] \quad (3)$$

to construct the Hessian matrix, $\mathbf{H}$.³⁸ Here, R_{ij}^0 and R_{ij} are the original and instantaneous distances between the C^α -atoms i and j , and Γ_{ij} is the ij th element of the Kirchhoff matrix $\mathbf{\Gamma}$, which is equal to -1 if $R_{ij}^0 < r_c$ and to zero otherwise, r_c is the cutoff-distance for inter-residue interactions, taken here as $r_c = 15$ Å. Eigenvalue decomposition of $\mathbf{H}$ yields $3N - 6$ nonzero eigenvalues, λ_k ($\lambda_1 \leq \lambda_2 \leq \dots < \lambda_{3N-6}$). The corresponding $3N$ -dimensional eigenvector $\mathbf{u}_k = (u_{k1}^x \ u_{k1}^y \ u_{k1}^z \ u_{k2}^x \ \dots \ u_{kN}^z)^T$ describes the distribution of the normalized collective displacements of the N residues along the k th mode, and $1/\lambda_k$ scales with the size of square displacements along that mode.

E. coMD simulations

coMD is a hybrid methodology that simulates conformational dynamics upon moving the structure collectively along the modes of motions predicted by the ANM. The moves, also called *steps*, are applied using a Monte Carlo/Metropolis algorithm. The steps are repeated until a predefined deformation, namely, cycle size, is attained. This defines an intermediate conformer, or *on-pathway conformer*, that is then targeted via tMD. The cycles are repeated until mapping the complete transition path. Details of the methodology can be found in our earlier study.²⁹

The transition OF $\leftrightarrow$ IF takes place either with two Na^+ ions and a substrate (Leu) bound (in the forward, OF $\rightarrow$ IF, direction), or in the apo form (in the IF $\rightarrow$ OF direction to complete the transport cycle). We examined the latter by coMD. For this purpose, we randomly selected OFo and IFo conformers from our unbiased MD simulations as our coMD starting points (e.g., snapshots 0.585 μs from run1 and 0.36 μs from run8a; see Table I); we removed the Na^+ ions from the selected OFo conformer and simulated the dynamics of

LeuT dimers (after energy minimization) using NAMD for 4 ns using stepwise reduced harmonic constraints on the C^α atoms (force constants of 4, 2, and 1 kcal mol⁻¹ Å⁻² were consecutively applied for 1, 1, and 2 ns, respectively). The conformation reached was then subjected to ANM analysis and deformed along the ANM modes selected by the coMD with a Metropolis acceptance ratio of 75%, i.e., 75% of energetically unfavorable motions (based on the deformation of a given protomer) were accepted (in addition to all favorable ones). A cycle size of 0.75 Å was adopted. The step size a was selected by setting the square displacement $a^2 \lambda_1^{-1} \mathbf{u}_1^T \mathbf{u}_1$ along the slowest ANM mode equal to 4 Å². Each tMD run of the coMD protocol was performed for 10 ps with a force constant of 20×10^3 kcal mol⁻¹ Å⁻². A total of 60 coMD runs were independently performed, each of which, on the average, consisted of 16 cycles, i.e., 32 coMD *on-pathway conformers* (shortly named *co-path conformer* x , with $1 \leq x \leq 32$) were sampled on the average per coMD, half of them starting from the OFo state and the other half, starting from the IFo state.

Conformers generated using the IFo form lacked the N-terminal segment R5-T10 (which was not resolved in the IFo crystal structure). In order to generate conformers with the N-terminal segment (under various binding states—apo/holo) and populate the regions of the conformational space where there were only very few conformers, we performed a 2nd round of simulations. This time, we started from *on-path* conformers (e.g., *co-path conformer* 27). To this aim, we examined all the conformers generated by cMD and identified among them snapshots in the *Ala/2Na⁺-bound*, *Leu/2Na⁺-bound* and *apo* states which are closest (in terms of RMSD) to the *co-path conformer* 27. These served as starting conformers in different (bound or unbound) forms to approach the vicinity of *co-path conformer* 27. To this aim we performed (for each case) a 2 ns tMD run, followed by 4 ns equilibration simulations. The equilibration simulations were conducted in two steps: 2 ns of MD in the presence of soft constraints applied to C^α -atoms ($k = 1$ kcal mol⁻¹ Å⁻²), and 2 ns without any constraints. The final conformations for each binding state were then used to guide the conformers to the next *co-path conformer*. By doing so, we populated the region between *co-path conformers* 27–30, which was not accessible to our original unbiased MD simulations. It should be noted that the transition between the IFo and OFo states, which are only 3.7 Å apart from each other (in terms of C^α RMSD), is spanned by 32 *co-path conformers*. This implies that the tMD simulations performed to move from one coMD intermediate to the other were very subtle.

III. RESULTS AND DISCUSSION

A. Energy landscape based on 21 μs cMD simulations shows multiple substates within the global states OF and IF of LeuT

In our recent study, we performed a large set of full-atomic MD simulations for the OFc and OFo LeuT dimers, including 16.5 μs trajectories generated for the WT transporter.²⁷ We extended those runs by an additional 2.4 μs in total, ending up with 18.9 μs of MD trajectories. In addition, we performed

2.1 μ s of MD simulations starting from the WT IFo crystal structure in LeuT dimer. A description of these runs is presented in Sec. II and in Table I.

To obtain the free energy landscape near the three structurally resolved states of LeuT, we first performed a PCA of the three (OFo, OFc, and IFo) crystal structures and determined the two principal axes that span the conformational space. Subsequently, MD snapshots were projected onto this principal frame. The occupancy of different grids on the surface, also called distributions of conformations, $f(\mathbf{R})$, was computed to estimate the free energy as $A(\mathbf{R}) = -kT \ln \{f(\mathbf{R})\} + \text{ct.}$ ³⁹ The free energy surface projected on the PCs 1-2 is shown in Fig. 1(a).

The free energy surface shows two distinctive energy wells, representative of the OF (*on the right*) and IF (*on the left*) states, each containing multiple minima (*substates*). The three crystal structures used as starting points are shown by the *gray squares*. The OF region connecting the OFo and OFc structures (see respective labels 3TT1 and 3F3E) was well sampled. These substates are separated by an energy barrier of ~ 2 kT. Simulations starting from the apo WT IFo structure (3TT3) also sampled more than one minimum. However, the region between the OF and IF states was not sampled by the simulations, mainly due to the inability of cMD runs to simulate the global transition between these two states in the absence of any bias. Toward a possible sampling of this region, we first generated a series of on-pathway conformers (shown by *black dots* in panel (a)), which were then used as initial conformers for performing additional cMD runs. These led to the energy landscape shown in Fig. 1(b), as explained in detail in Subsection III B.

B. New local minima emerge upon populating the rarely sampled regions of the conformational space with an additional 17.5 μ s of unbiased MD

As described in Sec. II, we performed 60 independent coMD runs in order to examine the transition pathway(s)

between the OFo and IFo forms of LeuT. The general features of these trajectories were very similar, if not identical, in support of the robustness of the results, and repeated runs helped improve our statistical analysis. Fig. 1(a) displays the projection of a representative pathway onto the free energy surface. We performed a new round of unbiased MD simulations, with a total duration of 17.5 μ s, starting from the *co-path* conformers occupying the scarcely populated regions of the space. These were performed under various binding conditions; apo, Ala/2Na⁺-bound, and Leu/2Na⁺-bound. For apo and Ala/2Na⁺-bound LeuT, the simulations were initiated from *co-path* conformers 29 and 30 based on their proximity to these forms; and for the Leu/2Na⁺-bound LeuT, *co-path* conformers 27 and 28 were selected as initial points (see Table I). This second set trajectories was combined with the first, which together provided $2 \times (17.5 + 21) = 77$ μ s long data for the WT LeuT monomer. The resulting new free energy surface is shown in Fig. 1(b). We note that the gap between the OF and IF states is narrowed down (compared to panel (a)). Furthermore, the new free energy surface indicates additional minima that were not sampled in the previous simulations, which have been labeled as OFc and *occl*, and IFc1 through 3.

In order to assess whether these new substates (or local minima) represent open or closed conformers, we reconstructed the free energy surface using the EC and IC vestibule openings as reaction coordinates. To this aim, we selected pairs of residues whose distances exhibit significant changes between the OF and IF states (Fig. 2) and used 2 pairs of inter-residue distances for quantifying the extent of opening of each of the two vestibules: $(k, l) = (\text{V33, D401})$ and (I245, I410) for the EC vestibule, and (R11, D369) and (I187, A268) for the IC vestibule. The inter-residue distance was expressed relative to its maximal value observed in simulations, as $s(k, l) = |\mathbf{R}_k - \mathbf{R}_l|/|\mathbf{R}_k - \mathbf{R}_l|_{\text{max}}$, and the average of $s(k, l)$ values over the two pairs of probes for each vestibule was evaluated as a quantitative measure of vestibule opening. The resulting energy maps, using the same

FIG. 1. Energy landscape for LeuT. (a) Two broad minima are observed, corresponding to the IF (*left*) and OF (*right*) states, and each state comprises local minima (e.g., OFc near PDB file 3F3E, and OFo near 3TT1 in the OF state). The surface is spanned by two principal components derived from the three LeuT crystal structures. The population/energy distribution is deduced from 21 μ s cMD generated for the dimer (yielding a total of 42 μ s trajectory, or 16 800 snapshots at 2.5 ns intervals, for the monomer). On-pathway conformations generated with coMD to explore the regions that could not be reached by cMD are shown in *black dots*. Starting conformers for coMD simulations (near 3TT3 and 3TT1) are shown by *black crosses*. (b) Same as in (a), obtained by combining two sets of simulations (a total of 30 800 snapshots or 38.5 μ s)—those starting from the crystal structures and a new set starting from coMD conformers labeled 27-30 and highlighted in white in panel (a). The surfaces are expressed relative to their lowest energy state in units of kT.

FIG. 2. Comparison of the EC and IC vestibule openings in the OFo ((a) and (c)) and IFo ((b) and (d)) of LeuT. Four residues are highlighted (shown in space-filling, colored) in each vestibule, which are used as metrics for probing the extent of exposure of the two vestibules (see text), used as axes in Fig. 3. Note the opening of TM1a (shown in red) in the IFo substate.

conformers as in Fig. 1, but now displayed as a function of these two new order parameters, EC vestibule opening and IC vestibule opening, are presented in Fig. 3.

Reconstruction of the energy map as a function of the extent of opening of the two vestibules showed that the *occl* region of Fig. 1(b) mapped to an *occluded-to-both-sides* (*occluded*) state in Fig. 3. Both the EC and IC vestibules are closed to the environment in this substate (see Fig. 3). The three new substates observed in the IFo state, labeled IFc1, IFc2, and IFc3 in Fig. 1(b), indeed represented IF substates where the IC vestibule was relatively closed. They are differentiated by the different levels of opening of the EC vestibule, while the IC vestibule remained relatively closed, IFc3 corresponding to the conformer with the largest opening at the EC vestibule.

In summary, based on these extensive simulations and analyses, we were able to identify three IF closed states (IFc1-3) and one substate occluded to both environments (*occl*) stabilized in the substrate-bound state (see below).

C. Some substates are preferentially assumed in the apo form, others in substrate/ion-bound state, thus defining a transport cycle path

It is important to note that the binding states of these minima were different. IFc1 as well as *occluded* states were sampled by substrate/ 2Na^+ -loaded form (with either Ala or Leu as substrate); whereas IFc2 and IFc3 were sampled by the apo transporter (Fig. 3 and Fig. S1⁴⁷). In addition to the Apo LeuT, the Ala/ 2Na^+ -loaded LeuT was also able to sample

IFc2. It is remarkable to observe that in the Leu/ 2Na^+ -loaded state the transition from OF to IF takes place through the *occluded* state, where both the EC and IC vestibules are fully closed, thus ensuring complete sealing of the cargo; whereas in the apo state, where the canonical transition back to the OF state takes place, both vestibules are less compact while maintaining relatively closed states. Ala-bound LeuT also stabilizes the occluded state, although the passage into IF state takes place through the IFc2 substate (where the EC vestibule is less tightly packed than the IFc1 state visited by Leu-bound LeuT). The arrows in Fig. 3(a) display the direction of the transport cycle during substrate/ion intake, translocation (*red arrows*), transition to IF state (*red* for Leu-bound LeuT, *green* for Ala-bound LeuT), release and back to uptake-ready state (*black*, apo form).

D. TM1a in the IFo state shows smaller opening in simulations relative to the crystal structure

The IFo crystal structure was obtained with 4 mutations: K288A, T354V, S355A, and Y268A. The latter three mutations are critical as they weaken the second sodium binding site and disrupt the cytoplasmic gate closure. In the IFo crystal structure, TM1a was tilted by $\sim 45^\circ$ with respect to its position in the OFc structure. The degree of this opening and its direction has been debated ever since, as it probably lifts TM1a into the membrane. Kazmier *et al.*⁴⁰ have recently performed measurements between spin-label pairs in LeuT and observed significantly smaller degree of openings of the IF vestibule for WT LeuT compared to the IFo crystal structure.

FIG. 3. Conformational substates and paths as a function of the degree of EC/IC vestibule openings. (a) The conformers in Fig. 1(b) are projected here onto the space spanned by the degree of opening of the EC (*abscissa*) and IC (*ordinate*) vestibules. ((b) and (c)) contributions to panel (a) by trajectories generated starting from Leu/2Na⁺-bound LeuT (b), Ala/2Na⁺-bound LeuT (c), and apo LeuT (d). The arrows in (a) display a tentative pathway for transport, starting from substrate/sodium uptake (OFo → OFc), closure to both media (*occl*), opening of IC vestibule (*occl* → IFc1 → IFo), release of substrate/ions and closure of IC vestibule, and return to OFo state (IFo → IFc1 → IFc2 → IFc3 → OFo) to start a new cycle. The intermediate substates sampled by Leu- and Ala-bound LeuT differ during the OF → IF transition, as distinguished by the *red* (for Leu-bound) and *green* (for Ala-bound) arrows.

However, in the Y268A and R5A mutants considerably larger openings were observed compared with WT LeuT. Their conclusion was that it is unlikely for TM1a to rotate all the way (45°) toward the middle of the membrane in the WT form and that the TM1a tilt in the crystal structure was likely due to the introduced mutations. In the current study, we have the WT LeuT protomer in the IF form during ~20 μ s of MD trajectory (out of 77 μ s) of which ~11.2 μ s correspond to the IFo state. Conformers having an IC vestibule opening larger than ~0.4 in Fig. 3 are identified as IFo substate, and those in the range [0.2-0.4] as IFc1-3. Our simulations exhibited striking results: IFo conformers in the WT LeuT showed smaller IC vestibule openings than the IFo crystal structure (the IC vestibule opening of which is ~0.7, indicated by the black square). Similarly, the OFc conformers showed slightly smaller EC vestibule opening than the OFc crystal.

Another metric of IC vestibule opening is the TM1a tilt angle with respect to its equilibrium position in the OF state. When we used a similar (if not identical) measure for TM1a opening as Krishnamurthy and Gouaux,²³ which is the angle of TM1a from its position in the OFc crystal structure, we observed a distribution with a peak at ~43° (Fig. 4), whereas that of the crystal structure was ~46°. These results are in accord with previous MD studies^{28,41} of several hundreds of ns of simulations for the apo and holo (Ala/2Na⁺-bound) K288A LeuT, showing slightly smaller TM1a openings than

the crystal structure. The smaller TM1a opening in our simulations also relates to the reorientation angle of TM1a, which differs from the one observed in the crystal structure; it opens rather toward TM7 which prevents it from lifting up into the membrane (Fig. 4(a)).

E. The K288A mutant samples the same conformational subspace as the WT LeuT

Many computational studies^{28,41,42} have focused on the dynamics of the K288A LeuT mutant instead of WT LeuT, mainly because this mutation has been shown to enhance substrate flux in proteoliposomes.⁴³ K288 is located on TM7 facing the membrane bilayer and distant from the LeuT central binding cavity. Whether this mutation causes a bias in the conformational space has not been investigated in detail. In order to address this question we have projected LeuT conformations obtained from 2.1 μ s of dimeric and 0.4 μ s of monomeric K288A-simulations onto the WT-LeuT free energy surface (shown in Fig. S2⁴⁷). Almost all K288A conformations fell into the regions which were previously sampled by WT LeuT *in silico*. This supports the use of the K288A structure as a biological model system which maintains equilibrium distribution of conformers while permitting us to visualize a higher number of transitions compared to the WT transporter.

FIG. 4. Analysis of the extent of tilting of the helical segment TM1a. (a) Representative IFo conformers from separate WT LeuT MD simulations. TM1a and TM7 are shown in *color*, whereas the remaining regions are shown in transparent. TM1a position in the OFc and IFo is shown in *red* for comparative purpose. (b) Histogram of TM1a rotations sampled in WT LeuT MD simulations. The inset shows the vicinity of the 2nd peak corresponding to the IF state. The red arrow indicates the $\sim 46^\circ$ TM1a tilt observed in the IFo crystal structure.

IV. CONCLUSION

In the present study, we thoroughly sampled the conformational space accessible to LeuT under native state conditions in order to gain insight into the energy landscape in the vicinity of the global OF and IF states and in the presence of Leu and Ala as substates. Data from multiple runs that yielded 77 μ s of unbiased trajectory have been compiled, which helped us map the probable directions of transition, or the paths undertaken by the transporter during its transport cycle. The sampling of the rarely visited regions was enhanced with the help of *coMD*. This method, building on our earlier hybrid methodology that utilizes ANM modes for guiding MD simulations,⁴⁴ permits to efficiently sample the conformational space between known states. The generated *coMD* trajectories indicated good agreement with the unbiased simulations starting from the crystal structures, the structures falling in many independent simulations into the same substates/minima that are preferentially selected in *cMD* runs. Our sampling suggests the EC vestibule and IC vestibule undergo sequential, almost independent movements during uptake and release (evidenced by the horizontal and vertical pathways in Fig. 3). We also examined the extent of tilting of TM1a, which has been a subject of interest in recent studies. In our simulations, TM1a undergoes a large tilt in the IFo state. However, it is not lifted up into the membrane as suggested by the IFo crystal structure (Fig. 4). The vestibule

opening in the IFo state in our simulations is smaller than that of the crystal structure. Similarly, the EC vestibule opening in the OFc state is slightly smaller than that in the crystal structure.

In our previous simulation of Ala transport by LeuT using advanced MD techniques,²⁸ two minimally hydrated intermediates, occluded to both EC and IC regions, were identified. The first stabilized by Ala-bound LeuT during the passage from outward-to inward-facing state (called *holo-occluded*²⁸) is in perfect agreement with the IFc2 state detected in the current unbiased simulations. The second, substrate-free, was sampled along the reverse transition (termed *apo-occluded*²⁸). In the current simulations multiple substates (IFc1, IFc2, and IFc3) are sampled during this reverse transition. These are relatively short-lived, and it is likely that the reverse transition indeed involves less well-defined intermediates (or multiple substates with different levels of EC vestibule opening, as the IC vestibule is relatively closed). The holo occluded intermediate²⁸ also overlap with *coMD* intermediates. While these substates have not been resolved in the LeuT family, the crystal structure of an occluded inward-facing state with bound Na^+ and substrate has been resolved for MhsT.⁴⁵ Strikingly, the resolved IFc structure shows close resemblance to the computationally predicted structure for human dopamine transporter,⁴⁶ suggesting that the passage over an occluded state is perhaps a common feature of LeuT-fold family members along their transport cycle.

ACKNOWLEDGMENTS

Anton computer time was provided by the National Center for Multiscale Modeling of Biological Systems (MMBioS) through Grant No. P41GM103712-S1 from the National Institutes of Health and the Pittsburgh Supercomputing Center (PSC). The Anton machine at MMBioS was generously made available by D. E. Shaw Research. I.B. gratefully acknowledges support from NIH Award Nos. P30 DA035778 and 5R01GM099738. M.G. thankfully acknowledges partial support from TUBITAK-BIDEB Award No. 115C038.

- ¹T. M. Jessell and E. R. Kandel, *Cell* **72**, 1 (1993).
- ²S. G. Amara and M. J. Kuhar, *Annu. Rev. Neurosci.* **16**, 73 (1993).
- ³L. L. Iversen and J. S. Kelly, *Biochem. Pharmacol.* **24**, 933 (1975).
- ⁴G. Rudnick, *Methods Enzymol.* **296**, 233 (1998).
- ⁵G. B. Richerson and Y. Wu, *Adv. Exp. Med. Biol.* **548**, 76 (2004).
- ⁶W. R. Schafer, *Cell* **98**, 551 (1999).
- ⁷J. Andersen, A. S. Kristensen, B. Bang-Andersen, and K. Stromgaard, *Chem. Commun.* **25**, 3677 (2009).
- ⁸R. P. Clausen, K. Madsen, O. M. Larsson, B. Frolund, P. Krogsgaard-Larsen, and A. Schousboe, *Adv. Pharmacol.* **54**, 265 (2006).
- ⁹M. K. Hahn and R. D. Blakely, *Pharmacogenomics J.* **2**, 217 (2002).
- ¹⁰I. D. Waldman, D. C. Rowe, A. Abramowitz, S. T. Kozel, J. H. Mohr, S. L. Sherman, H. H. Cleveland, M. L. Sanders, J. M. Gard, and C. Stever, *Am. J. Hum. Genet.* **63**, 1767 (1998).
- ¹¹J. R. Shannon, N. L. Flattem, J. Jordan, G. Jacob, B. K. Black, I. Biaggioni, R. D. Waldman, and D. Robertson, *N. Engl. J. Med.* **342**, 541 (2000).
- ¹²I. M. Anderson, *J. Affective Disord.* **58**, 19 (2000).
- ¹³O. Berton and E. J. Nestler, *Nat. Rev. Neurosci.* **7**, 137 (2006).
- ¹⁴A. Penmatsa, K. H. Wang, and E. Gouaux, *Nature* **503**, 85 (2013).
- ¹⁵G. Herting, J. Axelrod, and L. G. Whitby, *J. Pharmacol. Exp. Ther.* **134**, 146 (1961).
- ¹⁶R. W. Fuller, D. T. Wong, and D. W. Robertson, *Med. Res. Rev.* **11**, 17 (1991).
- ¹⁷S. G. Amara and M. S. Sonders, *Drug Alcohol Depend.* **51**, 87 (1998).

- ¹⁸H. Wei, E. R. Hill, and H. H. Gu, *Neuropharmacology* **56**, 399 (2009).
- ¹⁹A. Yamashita, S. K. Singh, T. Kawate, Y. Jin, and E. Gouaux, *Nature* **437**, 215 (2005).
- ²⁰O. Jardetzky, *Nature* **211**, 969 (1966).
- ²¹H. Krishnamurthy, C. L. Piscitelli, and E. Gouaux, *Nature* **459**, 347 (2009).
- ²²L. R. Forrest, Y. W. Zhang, M. T. Jacobs, J. Gesmonde, L. Xie, B. H. Honig, and G. Rudnick, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 10338 (2008).
- ²³H. Krishnamurthy and E. Gouaux, *Nature* **481**, 469 (2012).
- ²⁴S. K. Singh, C. L. Piscitelli, A. Yamashita, and E. Gouaux, *Science* **322**, 1655 (2008).
- ²⁵M. Quick, A. M. Winther, L. Shi, P. Nissen, H. Weinstein, and J. A. Javitch, *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5563 (2009).
- ²⁶S. K. Singh, A. Yamashita, and E. Gouaux, *Nature* **448**, 952 (2007).
- ²⁷E. Zomot, M. Gur, and I. Bahar, *J. Biol. Chem.* **290**, 544 (2015).
- ²⁸M. H. Cheng and I. Bahar, *PLoS Comput. Biol.* **10**, e1003879 (2014).
- ²⁹M. Gur, J. D. Madura, and I. Bahar, *Biophys. J.* **105**, 1643 (2013).
- ³⁰A. R. Atilgan, S. R. Durell, R. L. Jernigan, M. C. Demirel, O. Keskin, and I. Bahar, *Biophys. J.* **80**, 505 (2001).
- ³¹J. Henin, W. Shinoda, and M. L. Klein, *J. Phys. Chem. B* **112**, 7008 (2008).
- ³²M. Buck, S. Bouguet-Bonnet, R. W. Pastor, and A. D. MacKerell, Jr., *Biophys. J.* **90**, L36–L38 (2006).
- ³³R. B. Best, X. Zhu, J. Shim, P. E. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, Jr., *J. Chem. Theory Comput.* **8**, 3257 (2012).
- ³⁴J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, *J. Comput. Chem.* **26**, 1781 (2005).
- ³⁵D. E. Shaw, M. M. Deneroff, R. O. Dror, J. S. Kuskin, R. H. Larson, J. K. Salmon, C. Young, B. Batson, K. J. Bowers, J. C. Chao *et al.*, *Commun. ACM* **51**, 91 (2008).
- ³⁶W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graphics* **14**, 33 (1996).
- ³⁷J. Schlitter, M. Engels, and P. Kruger, *J. Mol. Graphics* **12**, 84 (1994).
- ³⁸Q. Cui and I. Bahar, *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems* (Chapman & Hall, 2005).
- ³⁹M. Gur, E. Zomot, and I. Bahar, *J. Chem. Phys.* **139**, 121912 (2013).
- ⁴⁰K. Kazmier, S. Sharma, M. Quick, S. M. Islam, B. Roux, H. Weinstein, J. A. Javitch, and H. S. Mchaourab, *Nat. Struct. Mol. Biol.* **21**, 472 (2014).
- ⁴¹J. Grouleff, S. Sondergaard, H. Koldso, and B. Schiott, *Biophys. J.* **108**, 1390 (2015).
- ⁴²M. H. Cheng and I. Bahar, *Biophys. J.* **105**, 630 (2013).
- ⁴³C. L. Piscitelli, H. Krishnamurthy, and E. Gouaux, *Nature* **468**, 1129 (2010).
- ⁴⁴B. Isin, K. Schulten, E. Tajkhorshid, and I. Bahar, *Biophys. J.* **95**, 789 (2008).
- ⁴⁵L. Malinauskaitė, M. Quick, L. Reinhard, J. A. Lyons, H. Yano, J. A. Javitch, and P. Nissen, *Nat. Struct. Mol. Biol.* **21**, 1006 (2014).
- ⁴⁶M. H. Cheng and I. Bahar, *Structure* **23**, 2171 (2015).
- ⁴⁷See supplementary material at <http://dx.doi.org/10.1063/1.4936133> for more details.