LETTER

doi:10.1038/nature09580

The mechanism of sodium and substrate release from the binding pocket of vSGLT

Akira Watanabe[1]*, Seungho Choe[2]*, Vincent Chaptal[1], John M. Rosenberg[2,3], Ernest M. Wright[1], Michael Grabe[2,3] & Jeff Abramson[1]

Membrane co-transport proteins that use a five-helix inverted repeat motif have recently emerged as one of the largest structural classes of secondary active transporters[1,2]. However, despite many structural advances there is no clear evidence of how ion and substrate transport are coupled. Here we report a comprehensive study of the sodium/galactose transporter from *Vibrio parahaemolyticus* (vSGLT), consisting of molecular dynamics simulations, biochemical characterization and a new crystal structure of the inward-open conformation at a resolution of 2.7 Å. Our data show that sodium exit causes a reorientation of transmembrane helix 1 that opens an inner gate required for substrate exit, and also triggers minor rigid-body movements in two sets of transmembrane helical bundles. This cascade of events, initiated by sodium release, ensures proper timing of ion and substrate release. Once set in motion, these molecular changes weaken substrate binding to the transporter and allow galactose readily to enter the intracellular space. Additionally, we identify an allosteric pathway between the sodium-binding sites, the unwound portion of transmembrane helix 1 and the substrate-binding site that is essential in the coupling of co-transport.

Secondary active transporters harness the energy stored in electrochemical gradients to drive the accumulation of specific solutes across cell membranes. This task is accomplished by the alternating-access mechanism, in which the substrate-binding site is first exposed to one side of the membrane and, on ion and substrate binding, a conformational change exposes the transported solute to the opposite face, where it is released[3]. Sodium/glucose co-transporters are prototypes of secondary active transporters that drive the accumulation of sugars and other molecules into cells. These transporters have critical roles in human physiology, where mutations in their genes are responsible for severe congenital diseases[4] and are the molecular targets for drugs to treat diabetes and obesity[5].

There has been a recent surge of work on crystal structures[6–11] displaying the five-helix inverted repeat motif. These are referred to as the LeuT superfamily and include genetically diverse proteins that transport a wide range of substrates and differ in the number and type of driving ligand[1,2,12]. A general model for alternating access is being pieced together through comparisons of these diverse structures[1,2,13]. Despite sharing a common set of ten core transmembrane segments, the lack of sequence similarity and the chemical diversity of the transported substrates prevents the complete understanding of the mechanistic basis of transport. This hurdle is being surmounted as multiple structures of the same protein—at different stages in the transport cycle—are solved, providing a comprehensive understanding of substrate binding[7,14–16] and the transition from outward- to inward-facing conformations[17]. However, an atomic-level understanding of sodium-coupled substrate co-transport, necessary to explain the dynamics of alternating access, is still absent.

To investigate the mechanism of sodium–sugar coupling, we carried out molecular dynamics simulations on the galactose-bound inward-occluded conformation of vSGLT[6] embedded in a lipid bilayer[18]. All sodium co-transporters of the LeuT superfamily share a common sodium-binding site termed the Na2 site. During the transition to the inward-facing conformation, transmembrane helix (TM) 8, which forms part of the sodium-binding site, is displaced by ~4 Å, generating a less favourable Na2 site that facilitates Na^+ release[1,17] (Fig. 1b). Na^+ modelled at this site is loosely coordinated by the carbonyl oxygens of Ile 65 (3.3 Å), Ala 361 (3.2 Å) and the side-chain hydroxyl of Ser 365 (3.1 Å). The carbonyl oxygen of Ala 62 (3.6 Å) and the side-chain hydroxyl of Ser 364 (3.6 Å) are also in close proximity (Fig. 1b). Previous molecular dynamics simulations performed on vSGLT[19] and Mhp1[17] indicated that Na^+ quickly leaves the Na2 site. Our simulations indicate that Na^+ exits the Na2 site after 9 ns (Fig. 2a) and interacts with the hydrophilic pore-lining residue Asp 189 on TM5 during exit. The importance of Asp 189 was highlighted in a previous simulation[19] and in biochemical studies on hSGLT1[20]. All three molecular dynamics simulations indicate that Na^+ exits the transporter before substrate exit; however, additional conformational changes are required to release the occluded galactose.

In the inward-occluded structure, galactose is located halfway across the membrane (Figs 1a and 2b), where it is coordinated by extensive side-chain interactions from TM1, TM2, TM6, TM7 and TM10. Subsets of these residues form two hydrophobic gates blocking galactose exit to the intracellular and extracellular spaces. Our molecular dynamics simulations show that as Na^+ exits the Na2 site, galactose undergoes significant fluctuations within the binding pocket. At 52 ns, Tyr 263 adopts a new and stable rotamer conformation that expands the exit pathway (between 52 and 110 ns), permitting the sugar to leave the binding site (Figs 2 and 3). After sugar release (~110 ns), Tyr 263 returns to the original conformation.

To test the hypothesis that Na^+ release stimulates an alternative conformation of Tyr 263, we conducted a 200-ns molecular dynamics simulation in which the sodium was lightly restrained in the Na2 site. Under these conditions, Tyr 263 never adopts the alternative conformation, and thereby prevents galactose exit (Supplementary Fig. 1). This observation suggests that sodium release drives conformational changes that disrupt the galactose-binding site, and further suggests that interactions between the Na2 site and Tyr 263 are central to the transport mechanism.

The spontaneous release of galactose in the absence of applied forces makes possible the accurate determination of the binding free energy profile through the use of umbrella sampling along the exit pathway coupled with weighted histogram analysis[21] (Fig. 3, inset). After Na^+ release, galactose is weakly bound to vSGLT with a minimal energy barrier of ~2 kcal mol^{-1}, resulting from the interaction of the sugar with residues Asn 64, Ser 66, Glu 68 and Gln 69 on TM1. Asn 64 is of particular interest because it is located in the unwound segment of TM1 and has hydrogen bonds with the inner gate residue Tyr 263 and the O2 hydroxyl of galactose linking the Na2 site with the galactose site. Thus, the interactions of Asn 64 with Tyr 263 and galactose may be critical to the transport mechanism[6].

To test the importance of these interactions, we performed molecular dynamics simulations and sodium-dependent transport assays on

[1]Department of Physiology, University of California, Los Angeles, Los Angeles, California 90095-1759, USA. [2]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA. [3]Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.
*These authors contributed equally to this work.

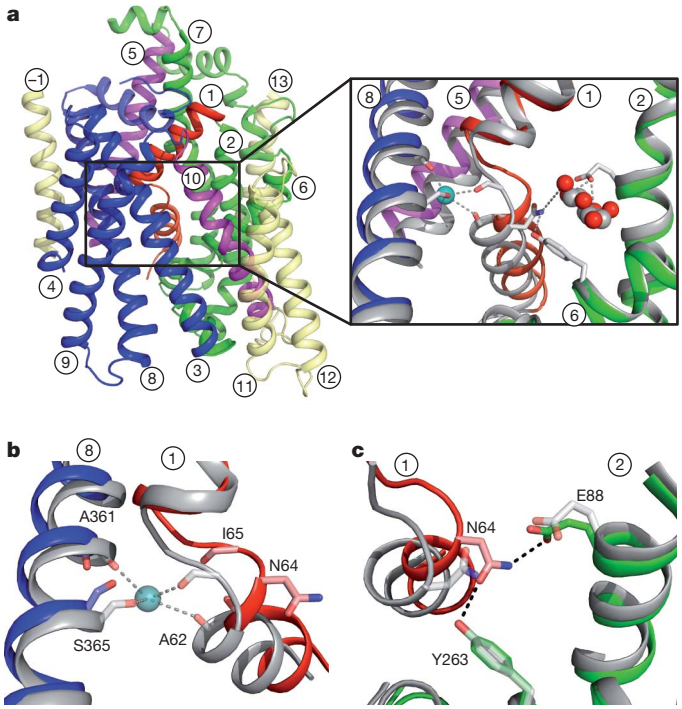

Figure 1 | Structures and overlay of the inward-open and inward-occluded conformations. **a**, The core domain of the inward-open conformation (TM1–TM10) is coloured by specific helix bundles involved in the transition from the inward-occluded to the inward-open conformation. The 'hash motif' formed from TM3, TM4, TM8 and TM9 is blue; the 'sugar bundle' formed from TM2, TM6 and TM7 is green; TM1 is red; and TM5 and TM10 are magenta. The periphery helices (TM−1, TM11, TM12 and TM13) are yellow. Atoms are displayed in ball-and-stick form with oxygen coloured red and nitrogen coloured blue. Inset, an overlay of the inward-open (colour) and inward-occluded (grey) conformations illustrating the coordination at the Na2 and galactose-binding sites. **b**, **c**, Overlay of the inward-open and inward-occluded conformations with the same colouring as in **a**. Conformational changes in the inward-open structure reveals a ~13° kink in the unwound segment of TM1 that prevents sodium coordination at the Na2 site (**b**). In the absence of galactose, the galactose-binding residue Asn 64 hydrogen-bonds to Glu 88 and Tyr 263, maintaining an open pathway from the intracellular space to the substrate-binding site (**c**).

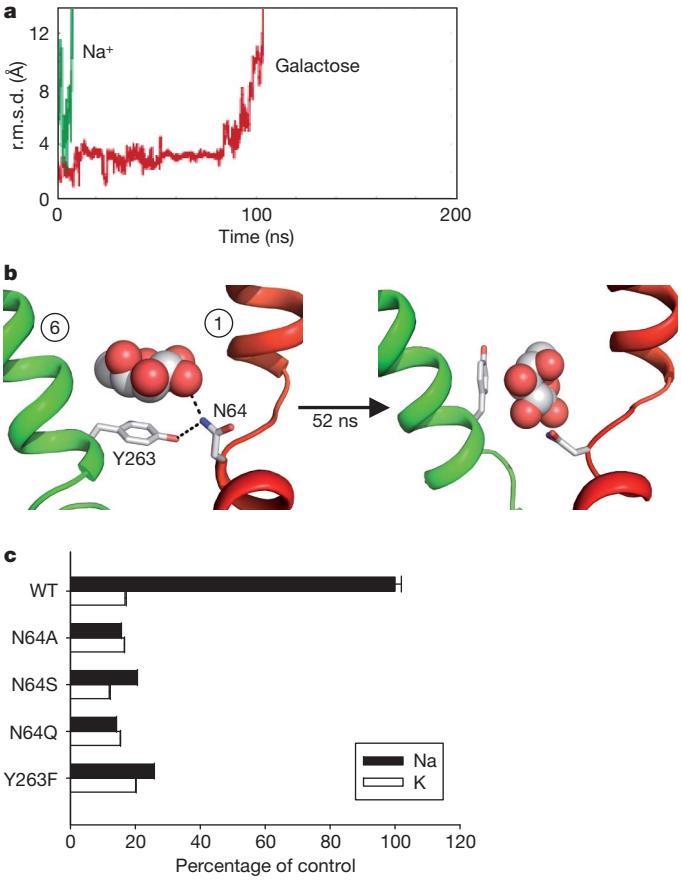

Figure 2 | Mechanism of galactose release. **a**, Sodium and galactose exit vSGLT. The root mean squared deviation (r.m.s.d.) of Na^+ (green) rapidly increases at 9 ns, indicating exit from the Na2 site. This is followed by the release of galactose (red) at 110 ns. **b**, Tyr 263 adopts two rotamers. On the left, Tyr 263 is shown in the conformation observed in the inward-occluded structure[6], in which it blocks substrate exit through a hydrogen bond with Asn 64 on TM1. At 52 ns (shown on the right), Tyr 263 adopts a rotamer conformation that expands the exit pathway. **c**, D-galactose uptake by wild-type and vSGLT mutants in proteoliposomes. Results are expressed as percentage uptake in either 100 mM NaCl or KCl, and show that the mutants Asn 64 Ala, Asn 64 Ser, Asn 64 Gln and Tyr 263 Phe severely impair sodium-dependent transport. Error bars, s.e.m. WT, wild type.

transporters with mutations at positions 64 and 263. Simulations of the Asn 64 Ala mutant show a momentary sodium departure from the Na2 site at 5 ns, but the ion rapidly returns and remains for the remainder of the simulation. The failure of Na^+ to unbind prevents conformational changes in the unwound segment of TM1, and Tyr 263 remains in the blocked orientation (Supplementary Fig. 2a). In agreement with the simulation, sodium-dependent transport assays on the Asn 64 Ala mutant show no activity (Fig. 2c).

To explore the role of Asn 64 further, we tested Asn 64 Gln and Asn 64 Ser, which, in principal, are both capable of maintaining the native hydrogen bonds to Tyr 263 and galactose. Models of Asn 64 Gln prevented simulation as the result of substantial steric clashes, which correlated well with a lack of transport (Fig. 2c). The model of the Asn 64 Ser mutation could form a hydrogen bond with Tyr 263 (3.3 Å) but not with the O2 hydroxyl of galactose (4.3 Å). In the simulation, Asn 64 Ser releases Na^+ from the Na2 site at 25 ns, and Tyr 263 transiently adopts the alternative rotamer conformation before returning to its original position, preventing galactose exit (Supplementary Fig. 2c). Similarly, simulation of the Tyr 263 Phe mutant shows that Na^+ unbinds at 10 ns, but unlike tyrosine, phenylalanine never adopts a conformation compatible with galactose exit (Supplementary Fig. 2b). Longer simulations may reveal galactose release in these mutants, because they both show modest transport activity (~10% of wild type); however, the transport assays and simulation data demonstrate that robust transport requires precise orientation of Asn 64 to stabilize galactose and the gating residue Tyr 263.

Although the molecular dynamics simulations and biochemical studies demonstrate a physical link between the Na2 site and the substrate, global details regarding the inward-open conformation (devoid of both ligands) remain elusive. To address this issue, we determined the structure of vSGLT in the inward-open conformation. Crystals, in the absence of ligands, for both the wild-type protein and the inactive Lys 294 Ala mutant[6] were obtained. Both crystals had the same overall configuration, but the mutant crystals diffracted to a higher resolution (2.7 Å; see Methods).

As in the original structure[6], the inward-open conformation is composed of 14 transmembrane helices, ten of which comprise the core domain. TM1–TM5 and TM6–TM10 are related by an approximate two-fold symmetry axis through the centre of the membrane plane. The inward-occluded and the inward-open structures have a similar overall fold with a r.m.s.d. of 1.2 Å. However, there are distinct structural differences between the two conformations, presumably owing to changes resulting from the release of ligands (Fig. 4). With the exception of TM1, superimpositions of individual helices reveal that the occluded-to-open transition occurs by rigid-body movements of subdomains (Supplementary Figs 3 and 4). Consistent with the recent

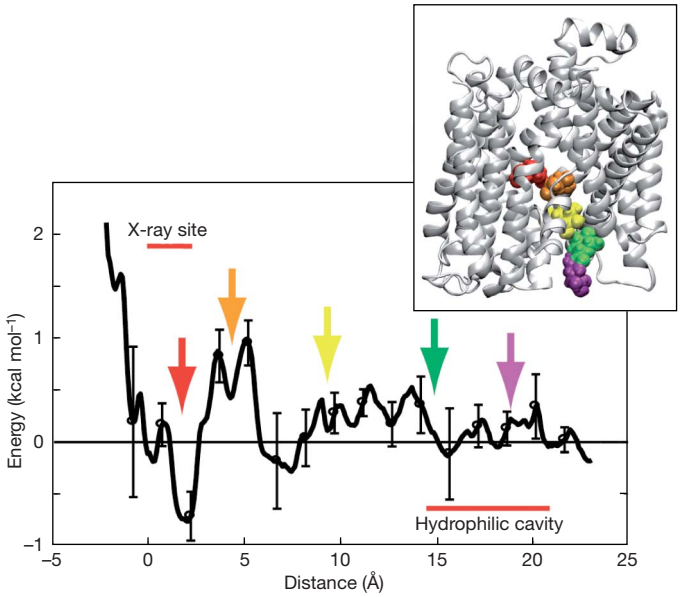

Figure 3 | The potential of mean force for galactose unbinding. Energy of galactose binding to vSGLT in the absence of Na^+. Umbrella sampling along the natural, equilibrium pathway shown (inset) was used to determine the binding free energy. The distance along the pathway from the binding site in the X-ray structure is shown along the *x* axis. The coloured arrows correspond to the galactose positions shown in the inset. The largest barrier is ~2 kcal mol^{-1}, at 5 Å, which corresponds to galactose interaction with residues in the kink region of TM1. Error bars were determined by splitting the production data into four equal sets, computing the energy profile for each set, and then applying a global shift to each curve before calculating the standard deviation at the 16 positions marked with points.

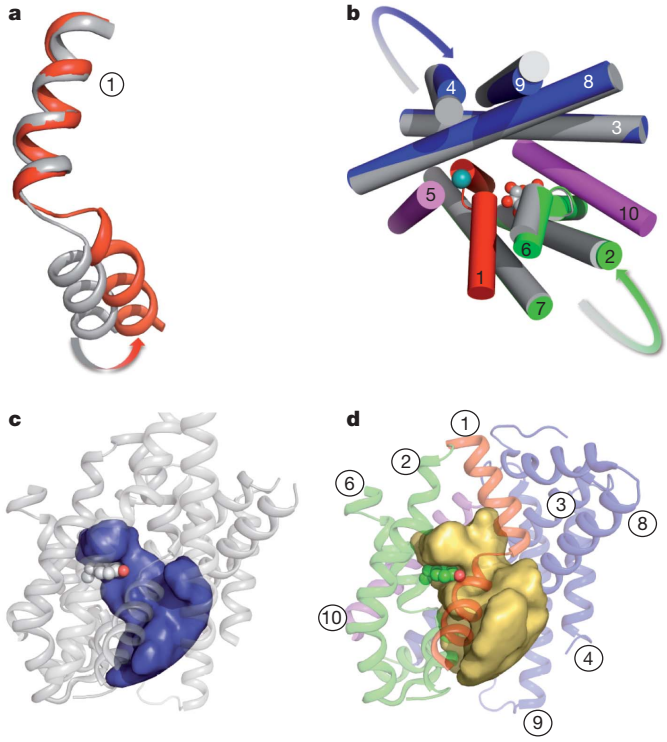

Figure 4 | Conformational changes in the transition from the inward-occluded to the inward-open structure. **a**, TM1 superimposed between the inward-open (red) and inward-occluded (grey) structures, showing a ~13° kink in TM1. **b**, Overlay of the inward-open (coloured as in Fig. 1) and inward-occluded (grey) conformations. Rigid-body rotations of the hash motif and sugar bundle by 3° in opposite directions expose the substrate-binding site to the intracellular environment. **c**, Accessibility cavity of the inward-occluded conformation is coloured blue. **d**, Accessibility cavity of the inward-open conformation is coloured gold. The conformational changes from TM1, hash motif and sugar bundle cause an increase of ~1,400 $Å^3$ in the accessible volume of the inward-open conformation, aiding galactose release.

assignment for Mhp1[17], the hash motifs, formed from TM3 and TM4 and their inverted repeat equivalents, TM8 and TM9, align with a r.m.s.d. of 0.9 Å. TM2, TM6 and TM7 form a domain termed the sugar bundle for the extensive side-chain interactions with galactose, and these regions superimpose with a r.m.s.d. of 0.5 Å. This new inward-open structure of vSGLT is more similar to the recent inward-facing conformation structure of Mhp1 than is the previous structure of vSGLT. Details of this structural analysis are in Methods.

The transition from the inward-occluded to the inward-open structure is presumably triggered by sodium release from the Na2 site and the alteration in the hydrogen-bonding network surrounding the unwound segment of TM1. In particular, the intracellular half of TM1 flexes ~13°, modifying the coordination of Asn 64 (Figs 1b and 4a). In the absence of both galactose and Na^+, Asn 64 coordinates Tyr 263 and Glu 88. Glu 88 was previously hydrogen-bonded to the O2 and O3 hydroxyls of galactose. This new conformation of TM1 is further stabilized by hydrogen bonds between the Na2-site residue Ser 365 and Glu 68 on the unwound segment of TM1 (Supplementary Fig. 5). When viewed from the intracellular side, each domain moves ~3° in opposite directions, thereby increasing the volume of the accessibility cavity by ~1,400 $Å^3$ (Fig. 4). This 6° relative rotation probably disrupts protein–substrate coordination and permits water to enter the site. This interpretation is supported by our simulations, in which an increase in the number of water molecules in the substrate-binding site is observed after sodium release (Supplementary Fig. 6). Water effectively competes with the protein for hydrogen bonds, loosening galactose in the pocket and ultimately assisting in its release (Supplementary Fig. 6 and Supplementary Movie 1).

We propose the following mechanism for sodium and galactose exit from vSGLT. The transition from the outward- to the inward-occluded conformation weakens the Na2 sodium-binding site, causing it to become metastable and release the ion on a short timescale. Upon exit, the carbonyl oxygens of the ion-coordinating residues Ile 65 and Ala 62 undergo a conformational change in the unwound segment of TM1, producing a kink of ~13° (Figs 1 and 4a). Our simulation shows that movement of TM1 disrupts the hydrogen bond between Asn 64 and Tyr 263, allowing the side chain of Tyr 263 to adopt a new conformation that opens a pathway to the intracellular space (Fig. 2b). Additional rigid-body movements widen the intracellular cavity, allowing water penetration and further disrupting the substrate-binding site to enhance exit and prevent rebinding (Fig. 4).

It is likely that the reaction scheme described here for vSGLT is broadly used by all sodium-dependent members of the LeuT superfamily, because the Na2 site, the hydrophobic gates and the unwound segments on TM1 and TM6 are all conserved[1,11,22,23]. For proteins with a single sodium-binding site, the Na2 site directly interacts with the substrate through polar residues—Asn 64 in vSGLT and Gln 42 in Mhp1—located on the unwound segment of TM1. For proteins that harbour two sodium-binding sites, such as LeuT and, putatively, BetP, the additional site (the Na1 site) is positioned on the opposite side of the unwound helix from the Na2 site. Interactions between the Na1 and Na2 sites are mediated by the unwound segment of TM1, and the sodium at the Na1 site is directly coordinated to the substrate[8,12]. Regardless of whether the protein has one or two sodium-binding sites, it is the conserved Na2 site, positioned most distal from the core of the protein, that regulates sodium and substrate release. This primary structural feature coupling sodium and substrate co-transport has fundamental implications for our understanding of membrane protein biology and for developing strategies to manipulate the alternating-access mechanism therapeutically.

METHODS SUMMARY

Molecular dynamics simulations. The vSGLT monomer (Protein Data Bank ID, 3DH4) was embedded and solvated in a 1-palmitoyl-2-oleoyl phosphatidylcholine membrane bilayer using the OPM[24] and CHARMM-GUI[18] software packages. Simulations were carried out using NAMD[25] with the CHARMM27 parameter set in a 150 mM NaCl bath. See Methods for more details.

Protein expression and purification. Plasmids carrying wild-type or mutant transporters were transformed and overexpressed in the TOP10 *Escherichia coli* cell line. Cell membranes were isolated, solubilized (2% w/v decyl-β-D-maltopyranoside) and tandem-purified using a Ni-NTA Superflow column (affinity chromatography) and a Superdex 200 column (size-exclusion chromatography). See Methods for more details.

Transport assays. We generated proteoliposomes by reconstituting purified vSGLT protein with sonicated lipid at a protein/lipid ratio of 1:200. We measured transport activity by monitoring the uptake of D-galactose, with ^{14}C-D-galactose tracer, into proteoliposomes in the presence or absence of a 100 mM Na^+ gradient (K^+ replacing Na^+). See Methods for more details.

Crystallization and data collection. We concentrated purified wild-type and Lys 294 Ala protein to $\sim$13 mg ml^{-1} and grew crystals by the hanging-drop vapour diffusion method using the Mosquito nanolitre-dispensing robot. Data collected at the Advanced Light Source, Berkeley (beamline 5.0.2), were integrated and scaled, and phases were calculated by molecular replacement. The model was built and refined to an R_{work}/R_{free} value of 25.1/27.4. See Methods for more details.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 May; accepted 12 October 2010.
Published online 5 December 2010.

1. Abramson, J. & Wright, E. M. Structure and function of Na^+-symporters with inverted repeats. *Curr. Opin. Struct. Biol.* **19,** 425–432 (2009).
2. Krishnamurthy, H., Piscitelli, C. L. & Gouaux, E. Unlocking the molecular secrets of sodium-coupled transporters. *Nature* **459,** 347–355 (2009).
3. Jardetzky, O. Simple allosteric model for membrane pumps. *Nature* **211,** 969–970 (1966).
4. Wright, E. M., Hirayama, B. A. & Loo, D. F. Active sugar transport in health and disease. *J. Intern. Med.* **261,** 32–43 (2007).
5. Isaji, M. Sodium-glucose cotransporter inhibitors for diabetes. *Curr. Opin. Investig. Drugs* **8,** 285–292 (2007).
6. Faham, S. *et al.* The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport. *Science* **321,** 810–814 (2008).
7. Weyand, S. *et al.* Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322,** 709–713 (2008).
8. Ressl, S., Terwisscha van Scheltinga, A. C., Vonrhein, C., Ott, V. & Ziegler, C. Molecular basis of transport and regulation in the Na^+/betaine symporter BetP. *Nature* **458,** 47–52 (2009).
9. Fang, Y. *et al.* Structure of a prokaryotic virtual proton pump at 3.2 Å resolution. *Nature* **460,** 1040–1043 (2009).
10. Gao, X. *et al.* Structure and mechanism of an amino acid antiporter. *Science* **324,** 1565–1568 (2009).
11. Shaffer, P. L., Goehring, A., Shankaranarayanan, A. & Gouaux, E. Structure and mechanism of a Na^+-independent amino acid transporter. *Science* **325,** 1010–1014 (2009).
12. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporters. *Nature* **437,** 215–223 (2005).
13. Forrest, L. R. & Rudnick, G. The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. *Physiology (Bethesda)* **24,** 377–386 (2009).
14. Zhou, Z. *et al.* LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. *Science* **317,** 1390–1393 (2007).
15. Gao, X. *et al.* Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* **463,** 828–832 (2010).
16. Singh, S. K., Piscitelli, C. L., Yamashita, A. & Gouaux, E. A competitive inhibitor traps LeuT in an open-to-out conformation. *Science* **322,** 1655–1661 (2008).
17. Shimamura, T. *et al.* Molecular basis of alternating access membrane transport by the sodium-hydantoin transporter Mhp1. *Science* **328,** 470–473 (2010).
18. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29,** 1859–1865 (2008).
19. Li, J. & Tajkhorshid, E. Ion-releasing state of a secondary membrane transporter. *Biophys. J.* **97,** L29–L31 (2009).
20. Quick, M., Loo, D. D. & Wright, E. M. Neutralization of a conserved amino acid residue in the human Na^+/glucose transporter (hSGLT1) generates a glucose-gated H^+ channel. *J. Biol. Chem.* **276,** 1728–1734 (2001).
21. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comput. Chem.* **13,** 1011–1021 (1992).
22. Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* **431,** 811–818 (2004).
23. Hunte, C. *et al.* Structure of a Na^+/H^+ antiporter and insights into mechanism of action and regulation by pH. *Nature* **435,** 1197–1202 (2005).
24. Lomize, A. L., Pogozheva, I. D., Lomize, M. A. & Mosberg, H. I. Positioning of proteins in membranes: a computational approach. *Protein Sci.* **15,** 1318–1333 (2006).
25. Kale, L. *et al.* NAMD2: Greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151,** 283–312 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Vondriska and K. Philipson as well as members of the Abramson, Wright and Grabe labs for useful discussions and for reading the manuscript. We would also like to thank S. Faham for contributions at the early stages of this work, S. Iwata for advance release of the Mhp1 coordinates (Protein Data Bank ID, 2X79), and R. Roskies for assistance with the computations. Simulations were carried out through a TeraGrid grant at the Pittsburgh Supercomputing Center and the Texas Advanced Computing Center. This work was supported by NIH grants GM078844 (J.A.), RGY0069 (J.A.) and DK19567 (E.M.W.), and a grant from the Human Frontier Science Program (J.A.). M.G. is an Alfred P. Sloan Research Fellow.

Author Contributions Experiments were carried out and diffraction data collected by A.W., V.C. and J.A. Simulations were carried out by S.C. Data were analysed and the manuscript was prepared by all authors.

Author Information Coordinates and structure factors of the inward-open vSGLT structure have been deposited in the Protein Data Bank under accession number 2XQ2. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.A. (jabramson@mednet.ucla.edu) or M.G. (mdgrabe@pitt.edu).

METHODS

Molecular dynamics simulations. Initially, TM−1 was removed and six missing residues in the TM4–TM5 loop were added with the loop modelling routine in Modeller[26]. Residues 53–547 were then embedded in a membrane and solvated in a hexagonal box approximately $96 \times 96 \times 84$ Å^3 in volume for a total of 63,000 atoms. Electroneutrality was enforced with the addition of 150 mM NaCl.

Simulations were carried out with the CMAP corrected[27] CHARMM27 parameter set and the TIP3P water model. VMD[28] and MATLAB were used for visualization and analysis. The system was minimized using conjugate gradient minimization and heated to 310 K using Langevin dynamics with a 10-ps^{-1} damping coefficient. An initial 300-ps equilibration using the NVT ensemble was carried out in which water, galactose, Na^+ and heavy backbone and side-chain atoms were constrained in a harmonic potential with a force constant of $k = 10.0$ kcal mol^{-1} Å^{-2}. We then switched to an NPT ensemble, and the restraints on the water molecules and heavy side-chain atoms were gradually removed in five steps over 1.5 ns. All remaining restraints were removed in six steps over the next 1.8 ns. Finally, 10 ns of restraint-free simulation was run. All production runs start from this equilibrated system. A Langevin piston with a 200-fs period and 100-fs decay was used to set the pressure to 1 atm. Hydrogen bond lengths were constrained with SHAKE[29], and a 2-fs time step was used. A 10 Å van der Waals cut-off was used along with the particle mesh Ewald method for the electrostatics.

Simulations with the restrained ion were carried out by holding the Na^+ in a weak spherical harmonic potential using the distance from the Na^+ to the centre of mass (COM) of the of the five coordinating residues. The equilibrium distance, based on the inward-occluded structure, was 1.39 Å, and a force constant of $k = 0.5$ kcal mol^{-1} Å^{-2} produced minimal distortions in the protein. A 10-ns equilibration was run before a 200-ns production run. The r.m.s.d. of the unrestrained simulation was 2.5 Å for the entire protein and 3.0 Å for the restrained-Na^+ simulation.

Potential of mean force calculation. The potential of mean force (PMF) was calculated using umbrella sampling with WHAM[21]. We extracted 69 snapshots along the pathway and held each configuration in a harmonic potential ($k = 7.0$ kcal mol^{-1} Å^{-2}) with a resting length equal to the z component of the distance between the galactose COM and the binding-site COM defined by the binding-site residues. Two nanosecond trajectories were run for each umbrella, and the last 1,800 ps were used for calculating the PMF. Splitting the trajectories into two equal parts (200–1,000 ps and 1,200–2,000 ps) and computing separate PMFs revealed that the total PMF is well converged.

Protein purification. vSGLT proteins were cloned, expressed and purified as previously described[6,30–32]. Briefly, the plasmids were transformed into the TOP10 cell line expressed to OD_{600} 1.8 and induced with 0.66 mM L-arabinose for 4 h at 29 °C. Cell membranes were isolated, solubilized with 2% decyl-β-D-maltopyranoside and affinity-purified on a Ni-NTA column. The sample was further purified by size-exclusion chromatography (Superdex 200) and washed with crystal buffer (20 mM Tris (pH 7.5), 25 mM NaCl, 0.174% decyl-β-D-maltopyranoside) in a 50-kDa Amicon filter unit.

Transport assays. Mutants were created with the QuikChange method and purified as above. vSGLT protein was reconstituted in 150 mM KCl, 10 mM Tris/Hepes (pH 8.0), 1 mM DTT, 1 mM Na_2EDTA, 1 mM $CaCl_2$, 1 mM $MgCl_2$ and 0.5% decyl-β-D-maltopyranoside, with 1.2 mg ml^{-1} sonicated lipid (90 mg asolectin soy lecithin, 10 mg cholesterol) at a protein/lipid ratio of 1:200. Addition of 5 mg ml^{-1} SM-2 Bio-Beads initiated the reconstitution and the mixture was incubated overnight at 4 °C. The proteoliposomes were collected and washed twice by centrifugation. Pelleted proteoliposomes were resuspended and underwent three freeze–thaw cycles in liquid nitrogen.

Uptake of D-galactose (88 μM) with ^{14}C-D-galactose tracer into proteoliposomes was measured for 18 min at 22 °C in the presence or absence of 100 mM Na^+ (K^+ replacing Na^+) as described previously[6,30]. Proteoliposomes were collected by filtration through 0.45-μm Millipore filters and the uptake was quantified by scintillation counting. Results are expressed as the mean ± s.e.m. of three determinations and three trials.

Crystallization. Protein was concentrated to ~13 mg ml^{-1} before plating. Optimization by additive screening gave the best diffracting crystals with a reservoir solution containing 0.1 M MES (pH 6.5), 4% MPD and 9–13% PEG400, and tridecyl-β-D-maltopyranoside to a final concentration of 0.0017% as an additive. Before freezing, crystals were cryoprotected using a solution containing 30% PEG400 and 0.174% decyl-β-D-maltopyranoside.

Data processing, phasing and refinement. Data was collected at 1.0 Å on cryo-cooled crystal (100 K) at the Advanced Light Source (beamline 5.0.2). Five data sets were integrated using HKL2000[33] and merged and subjected to B-factor-sharpening using an anisotropy correction server[34] (resolution cut-offs: $a = 3.1$ Å, $b = 2.7$ Å and $c = 2.8$ Å). Phases were calculated by molecular replacement (PHASER[35]) using the original vSGLT structure as a search model. The model was built in COOT[36] and refined using PHENIX[37] and BUSTER[38] using non-crystallographic symmetry (NCS) and TLS refinement restraints. There are two molecules per asymmetric unit with the A molecule displaying sharper electron density and lower B factors (88.5 Å^2) than the B molecule (131.3 Å^2). The model was built and refined to an R_{work}/R_{free} value of 25.1/27.4. The Ramachandran statistics shown areas follows: 95.5% of the residues lie in the preferred region, 4.3% lie in the allowed region and 0.2% are outliers.

The $2F_o–F_c$ maps contained three elongated features having a maximal peak height of 3σ. These attributes were interpreted and assigned as PEG molecules. Two are located at the periphery, whereas the third is near the Na2 site as observed in the Mhp1 structure[17] and is proposed to stabilize the inward-facing conformation.

The Lys 294 Ala protein crystals diffract to higher resolution than the wild-type crystals. Data from four wild-type crystals were collected and merged to achieve a 3.7 Å resolution data set. Difference Fourier maps were calculated against the final Lys 294 Ala mutant model and no significant peaks were observed. The Lys 294 Ala model was further refined using PHENIX to yield an R_{work}/R_{free} value of 30.7/34.8.

We note that although refinement was carried out with data subject to anisotropic correction, as described above, the deposited data has not been treated. Figures were created from the A-chain protomer using PYMOL[39].

Structural comparison of vSGLT with Mhp1. Superpositions of the inward-occluded (Protein Data Bank ID, 3DH4) and inward-open conformations of vSGLT with the inward-facing conformation of Mhp1 (Protein Data Bank ID, 2X79) reveal they all share a similar global fold. The largest differences are centred near the substrate- and ion-binding sites. The Na2-site helices (TM1, TM5 and TM8) of the inward-open conformation of vSGLT have a closer fit to Mhp1 (r.m.s.d., 2.2 Å) than the inward-occluded conformation (r.m.s.d., 2.6 Å); thus, the inward-open vSGLT structure more closely resembles the structure of Mhp1.

26. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815 (1993).
27. MacKerell, A. D. Jr, Feig, M. & Brooks, C. L. III. Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **126,** 698–699 (2004).
28. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38 (1996).
29. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23,** 327–341 (1977).
30. Turk, E. *et al.* Molecular characterization of Vibrio parahaemolyticus vSGLT: a model for sodium-coupled sugar cotransporters. *J. Biol. Chem.* **275,** 25711–25716 (2000).
31. Turk, E., Gasymov, O. K., Lanza, S., Horwitz, J. & Wright, E. M. A reinvestigation of the secondary structure of functionally active vSGLT, the Vibrio sodium/galactose cotransporter. *Biochemistry* **45,** 1470–1479 (2006).
32. Veenstra, M., Lanza, S., Hirayama, B. A., Turk, E. & Wright, E. M. Local conformational changes in the Vibrio Na^+/galactose cotransporter. *Biochemistry* **43,** 3620–3627 (2004).
33. Otwinowski, Z. & Minor, W. in *Macromolecular Crystallography* (eds Abelson, J. N., Simon, M. I., Carter, C. W. Jr & Sweet, R. M.) 307–326 (Methods in Enzymology 276, Academic, 1997).
34. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proc. Natl Acad. Sci. USA* **103,** 8060–8065 (2006).
35. Qian, B. *et al.* High-resolution structure prediction and the crystallographic phase problem. *Nature* **450,** 259–264 (2007).
36. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
37. Adams, P. D. *et al.* Recent developments in the PHENIX software for automated crystallographic structure determination. *J. Synchrotron Radiat.* **11,** 53–55 (2004).
38. Joosten, R. P., Womack, T., Vriend, G. & Bricogne, G. Re-refinement from deposited X-ray data can deliver improved models for most PDB entries. *Acta Crystallogr. D* **65,** 176–185 (2009).
39. DeLano, W. L. *PyMOL Molecular Viewer* ⟨http://www.pymol.org⟩ (2002).