Published in final edited form as:
Science. 2008 August 8; 321(5890): 810–814. doi:10.1126/science.1160406.

The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/sugar symport

Salem Faham[1], **Akira Watanabe**[1], **Gabriel Mercado Besserer**[1], **Duilio Cascio**[2], **Alexandre Specht**[3], **Bruce A. Hirayama**[1], **Ernest M. Wright**[1,*], and **Jeff Abramson**[1,*]

[1]Department of Physiology, David Geffen School of Medicine, University of California, Los Angeles, California 90095-1751, USA

[2]UCLA-Department of Energy Institute of Genomics and Proteomics, University of California, Los Angeles, CA 90095, USA

[3]Laboratoire de chimie bioorganique, Université Louis Pasteur / CNRS UMR 7175 LC01, Faculté de Pharmacie, 74 route du Rhin, 67401 Illkirch, France

Abstract

Membrane transporters that use energy stored in sodium gradients to drive nutrients into cells constitute a major class of proteins. We report the crystal structure of a member of the solute sodium symporters (SSS), the *Vibrio parahaemolyticus* sodium/galactose symporter (vSGLT). The ~3.0Å structure contains 14 transmembrane helices in an inward facing conformation with a core structure of inverted repeats of 5 TM helices (TM2-TM6 and TM7-TM11). Galactose is bound in the center of the core, occluded from the outside solutions by hydrophobic residues. Surprisingly, the architecture of the core is similar to the leucine transporter (LeuT) from a different gene family. Modeling the outward-facing conformation based on the LeuT structure, in conjunction with biophysical data, provides insight into structural rearrangements for active transport.

A central question in biology is how energy is harnessed to do work. For the active accumulation of glucose into cells, Crane proposed that energy was obtained from the inward Na^+ gradient, i.e. Na^+/glucose cotransport (symport) (1). This hypothesis has been extensively tested, refined and expanded to include the active transport of solutes and ions in virtually all cell types (2). It is now established that cells maintain a low intracellular [Na^+] through the active pumping of Na^+ out of the cell. This inward Na^+ gradient, along with a negative membrane potential, drives the transport of substrates into cells. Despite this paradigm, the structural basis for Na^+ solute symport is unknown.

Solute Sodium Symporters (SSS) (TC# 2.A.21) are a large family of proteins that cotransport Na^+ with sugars, amino acids, inorganic ions or vitamins (3). Members of this family are important in human physiology and disease where mutations in glucose and iodide symporters (SGLT1 and NIS) result in the congenital metabolic disorders glucose-galactose-malabsoption (GGM) and iodide transport defect (ITD) (4, 5). SGLT1 is the rationale for oral rehydration therapy and SGLTs are currently being targeted in drug trials for type II diabetes.

The first member of the SSS family to be cloned was the intestinal Na^+/glucose symporter (SGLT1) (6) and since then over 250 other members have been identified across all six kingdoms of life. The functions of the family members have been well characterized, including the mammalian glucose (SGLTs), the iodide (NIS), the *Vibrio parahaemolyticus*

To whom correspondence may be addressed. ewright@mednet.ucla.edu; jabramson@mednet.ucla.edu.

galactose/glucose (vSGLT) and the *Escherichia coli* proline (PutP) symporters (3, 7, 8). vSGLT has a sequence identity of 32% (60% similarity), 19% (58% similarity) and 18% (57% similarity) to SGLT1, NIS and PutP (SFigure 1). All share an alternating-access mechanism with tight coupling between Na^+ and solute transport (9–11). Despite functional studies supporting multiple states for symport, an absence of structural data on the protein family has precluded rigorous examination of this hypothesis.

To gain structural insight into the mechanistic details, we solved the structure of vSGLT in the presence of Na^+ and galactose. As predicted by experimental and *in silico* studies, vSGLT has 14 membrane spanning helices (TM) with extracellular amino- and carboxy-termini (12, 13). The structural core is formed from inverted repeats of 5 TM helices (TM2-TM6 and TM7-TM11) placing it in an inward facing conformation. Galactose is bound in the center of the core, occluded from the outside solutions by hydrophobic residues. We use structural and functional data, coupled with modeling, to predict the conformational changes as the transporter alternates between the outward and inward facing conformations, thereby providing a structural explanation for symporter dynamics.

Structure determination

Optimized crystals of SGLT from *Vibrio parahaemolyticus* displayed anisotropic diffraction, with Bragg spacing of 2.7 Å and 3.7 Å along the best and worst directions, respectively. Data were collected for 3 different crystal forms: P1, $P2_1$, and $P2_12_12_1$. Initial phases were obtained in the $P2_1$ crystal form, permitting the building of a partial model that we used for molecular replacement into the P1 and $P2_12_12_1$ crystal forms. Multi-crystal averaging and solvent flattening substantially improved experimental phases and Se-methionine crystals were used for sequence assignment. The model was refined from merged data to a resolution of 2.7 Å with an R-free of 29.5% and an R-work of 26.9% (STable 1). Of the 548 residues of vSGLT, 513 were built; the remaining residues are in disordered loop regions (14).

Transporter architecture

The protein assembles as a tightly packed parallel dimer with an extensive contact surface area of 1488$Å^2$ occurring between TM helices within the lipid bilayer (SFigure 2). The functional relevance of the dimer is unclear, as freeze-fracture electron microscopic studies indicate that SGLTs function as monomers (13, 15). The overall architecture has a maximal height and diameter of ~75 Å and ~55 Å (Figure 1). There are 14 transmembrane helical segments (TM1 to TM14) with both N- and C-termini exposed to the periplasm. A single galactose molecule is bound in the center of the protomer. A structural characteristic, not anticipated from amino acid sequence, is the inverted topology repeat motif; TM2-TM6 and TM7-TM11 are related by a ~153° rotation parallel to the membrane plane and are structurally related (Figure 1A and SFigure 4A,B). These ten transmembrane helices likely represent the core structure in the SSS family, with additional helices located on the N- and C-termini. This inverted topology repeat appears to be a common motif of membrane transport proteins (16–20).

The structure is comprised of a central group of 7 helices (TM2, TM3, TM4, TM7, TM8, TM9 and TM11) that contribute side chain interactions for ligand selectivity, along with 7 supporting helices that stabilize these central helices (Figure 1 and SFigure 3C). The positioning of the central helices in complex with the substrate forms a large cavity that extends from just below the galactose-binding site to the intracellular space, representing the inward-facing conformation (Figure 4B). A striking feature is the two discontinuous transmembrane helices, TM2 and the symmetrically related TM7, in the center of the protomer. In TM2 there is a break in the hydrogen-bonding pattern around residues I65, S66

and A67 dividing it into roughly equivalent intracellular (TM2I) and extracellular (TM2E) components. In TM7 the helix is disrupted at residues F266, N267, Q268 and Y269, dividing it into a shorter intracellular (TM7I) and a larger extracellular (TM7E) segment (Figure 1, 2A). This structural feature may have particular functional significance, as discontinuous helices have been implicated in transport mechanisms of several cotransporters (18, 20, 21).

There are notable structural features in the solvent-accessible regions. On the intracellular side, helix IL3 situated between TM3 and TM4 (Figure 1), is located on the outer edge of the cavity leading to the substrate-binding site. On the extracellular face between TM8 and TM9 there are two extracellular helices (EL8Ha) and (EL8Hb). There may be mechanistic implications with EL8Hb as it straddles the membrane plane and forms extensive contacts with the central helices TM2 and TM4. There is an additional extracellular helix (ELH6) between TM6 and TM7. This loop connects the two inverted repeats and harbors a glycosylation site in the eukaryotic SGLTs (22).

The 10 core helices (TM2-TM11) form a structure that displays a similar topology to the core structure of LeuT (TM1-TM10), a bacterial member of the Neurotransmitter Sodium Symporters (NSS) family (18) (SFigure 4C). This structural homology was surprising, because the SSS and NSS families have no significant similarity at the sequence level. Sequence based programs were unable to find the alignment between LeuT and vSGLT, and a structure based alignment only displayed an 11.5% identity. These findings support classification of proteins using criteria such as topological arrangement, molecular function and unique structural features involved in mechanism, rather than solely on the basis of primary sequence (23).

Galactose-binding site

Galactose analogs modified with electron dense atoms were used to identify the precise position and orientation of the sugar molecule. Galactose is bound approximately halfway across the membrane bilayer by specific side-chain interactions from the central helices TM2E, TM3, TM7E, TM8 and TM11 (Figure 2). The galactose-binding site is sandwiched between hydrophobic residues that form intracellular and extracellular gates. On the intracellular side, Y263 from the discontinuous helix TM7E stacks with the pyranose ring, a common feature seen in all sugar-binding structures to date (24, 25). This primary interaction, along with the flanking residues Y262 and W264, establishes a gate that prevents exit of the sugar to the large hydrophilic cavity contiguous with the intracellular compartment. The extracellular gate is formed by a triad of hydrophobic residues (M73, Y87 and F424) (Figure 2B). Directly above the hydrophobic residues, there are substantial interactions between TM11, TM3, TM2E, and TM7E with the loops from TM2-TM3, TM8-TM9 and TM10-11. Taken together, these interactions comprise considerable protein mass that extends from the substrate-binding site to the extracellular surface, thereby further blocking access of the sugar-binding region from the extracellular milieu. In the members of the SSS family that bind sugars, the residues forming the hydrophobic gates are highly conserved (SFigure1).

All OH-groups of the galactose ring are coordinated through H-bonds (Figure 2C). A possible H-bond between the carbonyl oxygen of Y87 and the C4-OH of galactose is also observed. Because Y87 is part of the extracellular hydrophobic gate, this additional H-bond may have mechanistic implications for galactose binding and stabilization in the central domain. K294 forms a H-bond with the C2-OH in a coordination similar to that observed in the glucose/galactose binding protein (PDB: 1GCA) and lactose permease (PDB: 1PV7); indeed, H-bonding with a positively charged side chain is a common feature of sugar

binding proteins (24–26). Overall, vSGLT supports the sugar by extensive side-chain interactions from the central domain and held in an occluded conformation through intracellular and extracellular hydrophobic gates.

To verify the structural data and understand the functional significance of substrate binding, residues forming the sugar-binding site were mutated to alanine and galactose transport was measured in proteoliposomes (Figure 2D). The Q69A, E88A, N260A, K294A and Q428A mutations did not show Na^+-dependent galactose transport. In contrast, the S91A mutant behaved normally, suggesting that it is not essential for protein function (S91 is not conserved across different species). Mutants of the corresponding sugar-binding site residues in SGLT1 (K321A, Q457R and Q457C: SFigure 1A) provide strong support for the argument that these residues also play an important role in the mammalian transporter (27, 28). The mutant Q457R causes severe glucose-galactose malabsorption in human subjects, a potentially fatal disease in newborn infants (4). This indicates a conserved role for these sugar-coordinating residues in protein function across isoforms and species.

Sodium-binding site

Although the Na^+ electrochemical gradient is the driving force for symport, Na^+/substrate stoichiometry varies amongst SSS members; for SGLT1 and NIS the stoichiometry is 2:1 (10, 11), while for vSGLT and PutP, it is 1:1 (8, 13). Nevertheless, it is expected that the mechanisms that couple transport of Na^+ and substrate will have commonality within the SSS family.

As in other transporters, precise localization of the Na^+ binding site(s) in vSGLT is not easily identified due to the similarity in density between Na^+ and water. Co-crystallization was attempted with electron-dense ions (Rb, Cs, and Tl), but none were observed in the resulting structure. Nonetheless, we were able to identify a plausible Na^+-binding site based on a comparison with the LeuT structure (18), conservation of sequence amongst SSS proteins and a mutational analysis. Structural alignment with LeuT revealed a possible Na^+-binding site at the intersection of TM2 and TM9, approximately 10Å away from the substrate-binding site (Figure 3). In this region, the carbonyl oxygens from A62 (TM2E), I65 (discontinuous segment of TM2) and A361 (TM9) would be appropriately positioned to coordinate the Na^+. An interaction is also possible with the hydroxyl oxygen of S365 of TM9, a residue conserved throughout the SSS family (SFigure 1). This site has a similar proximity as well as coordination properties to the second Na^+ site proposed for LeuT (18): Superimposition of the two structures reveals localization of these Na^+ sites in the two protein structures within ~2Å (Figure 3B). Functional studies provide supporting evidence for Na^+ binding in this region. In PutP, a hydroxyl moiety at position T341 is essential for Na^+ binding (29), and studies on NIS have shown that T354 plays a key role in Na^+ -dependent transport of iodide (30). Both of these residues are at equivalent position to S365 on TM9 in vSGLT. Mutation S365A in vSGLT completely abrogated Na^+ -dependent transport (Figure 2D). Unlike LeuT, but consistent with biochemical data, we did not detect a second Na^+ binding site in vSGLT.

The inward facing conformation

vSGLT resides in an inward facing conformation with the galactose-binding site inaccessible to the extracellular and intracellular compartments (Figure 2). The intracellular exit pathway appears as a large hydrophilic cavity blocked by the intracellular gate residue (Y263 on TM7E; Figure 2C). The cavity is formed from the intracellular portions of helices TM2I, TM3, TM4, TM7I, TM9 and TM11 (Figure 4). This finding is consistent with accessibility studies on PutP (29). Cysteine scanning mutagenesis on TM2 and TM9 of PutP reveals a hydrophilic face accessible to bulky sulfhydryl reagents (fluorescein-5-maleimide),

indicating an aqueous pathway between the site of substrate-binding and the cytoplasm (29, 31). For NIS, five hydroxyl-containing residues (T351, S353, T354, S356, and T357) and N360 located along one face of TM9 play a key role in Na^+/iodide cotransport (30). Together, these data suggests that the inward-facing conformation of vSGLT hosts a large cavity exposed to the cytoplasm that requires simple displacement of an intracellular gating residue (Y263) for release of galactose.

Alternating access

According to kinetic models for Na^+/solute symport (8–10) the proteins alternate between outward and inward facing conformations. Thus, for galactose to enter the vSGLT binding site from the extracellular side, it must adopt an outward facing conformation. Clues about this conformation arise from the structure of LeuT (18) which is in an outward-facing conformation. Given the structural similarity between vSGLT and LeuT, we mapped the core domain of vSGLT onto that of LeuT using the program MODELLER (32). Comparison of the two conformations (Figure 4 A & B) in conjunction with biochemical data indicates a series of specific helical rearrangements and a plausible pathway for Na^+ and galactose entry. Specifically, modeling reveals an external pathway to the substrate-binding site formed by TM2E, TM3, TM7E, TM11 and helix EL8b in the extracellular loop. External Na^+ binds first (33), presumably at the site identified in Figure 3 and we postulate that facilitates molecular rearrangements in TM2 to form the substrate-binding site. Galactose binding will induce the formation of extracellular gate (Y87, F424 and M73) closing the cavity through bends in TM3 and TM11. These structural rearrangements are facilitated by conserved glycine and proline residues (TM3 G99 and P104; TM11 P436 and G437). Consistent with this model, the corresponding helices of LeuT have conserved glycine residues in the same regions. Further evidence for these structural rearrangements is corroborated by site-specific fluorescence experiments on vSGLT, hSGLT and PutP (8, 33–35).

Larger conformational change occurs in the intracellular portions of transmembrane segments TM3-TM4, TM7-TM8 and TM9-TM10-TM11 to expand the intracellular cavity. This places the protein in the state represented by our structural data: inward-facing but blocked by the internal hydrophobic gate (Y263). Although the residency time in the occluded state is unknown, the protein must next shift conformation to release Na^+ and then galactose by displacing the intracellular gate (i.e. displacement of Y263). Supporting this mechanism are studies on hSGLT that document a conformational change in the protein following sugar binding (9), which may be indicative of either occlusion of the external cavity and/or exposure of the internal cavity.

The structure of vSGLT in an inward facing conformation, along with the structure of the functional and structural homolog LeuT, provides a model of the conformational changes underlying the alternating access mechanism. These structural and functional similarities indicate a common transport mechanism between members of the SSS and NSS gene families. To expand on this transport mechanism, further structural and functional analyses are required. Finally, given the functional homology between the bacterial and humans SGLTs, the vSGLT structure should provide a critical tool in the rational design of SGLT drugs for the treatment of diabetes.

Acknowledgments

The authors thank Eric Turk (contributions to the early stages of the project); Julian Whitelegge (assistance with the ESI-MS spectroscopy); and Ronald Kaback, Ken Philipson, James Bowie, Thomas Vondriska, Vincent Chaptal, Don Loo (critical comments on the manuscript). Michael Hahn and Daniel Kwon helped in sample preparation and the Protein Expression Core made large-scale protein production possible. The Advance Light Source (ALS),

Advance Photon Source and Swiss Light Source synchrotron were all used. The authors would like to emphasize the efforts of Peter Zwart and the beamline staff at 5.0.2. The coordinates are deposited at the PDB accession code: 3DH4. Supported by NIH grants DK19567 (EMW), DK44602 (EMW) and NIH GM07844 (JA), HFSP RGY0069 (JA) and AHA 0630258N (JA).

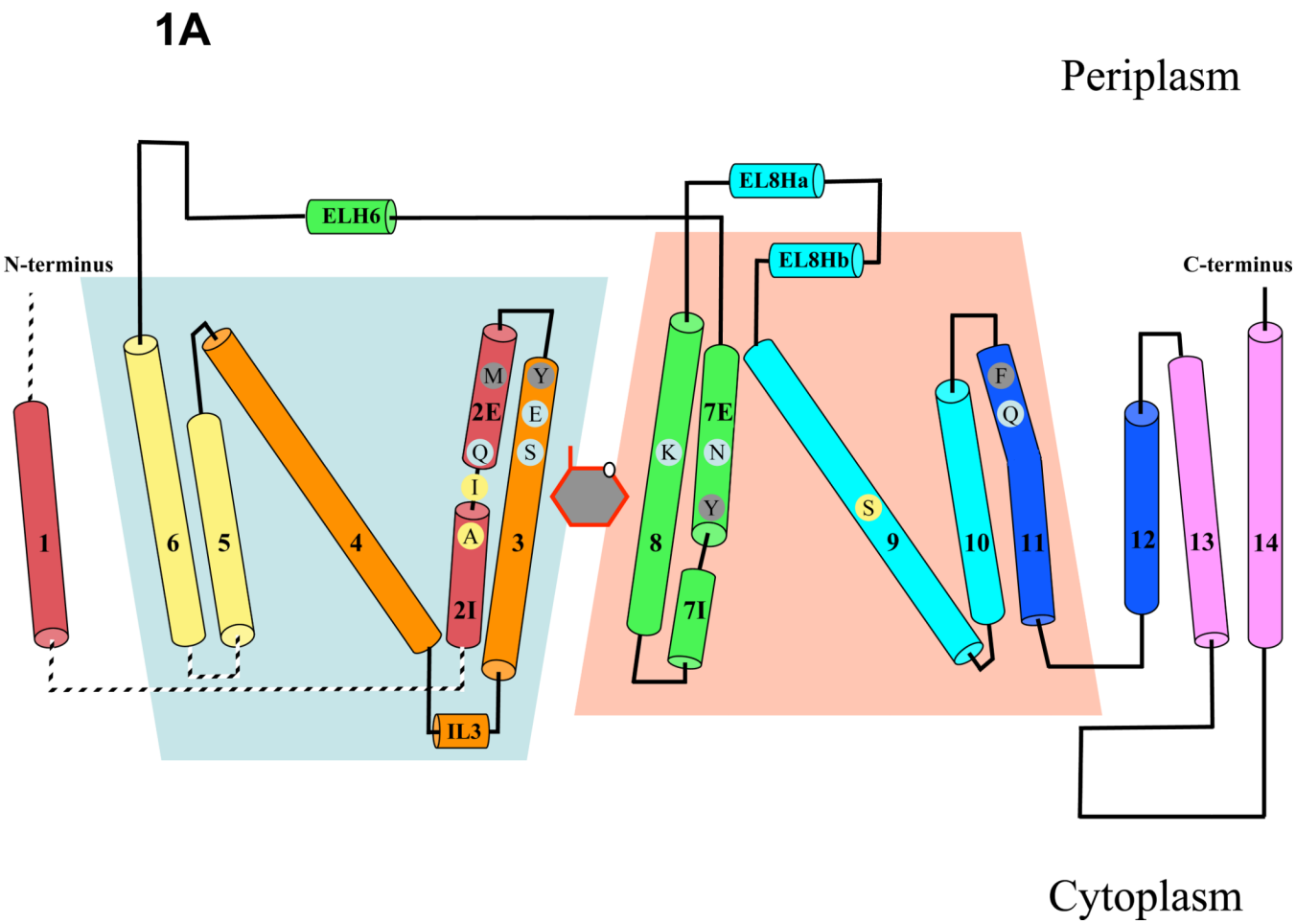
References

1. Crane, R.; Miller, D.; Bihler, I. Membrane Transport and Metabolism. Kleinzeller, A.; Kotyk, A., editors. London: Academic Press; 1961. p. 439-449.
2. Schultz SG, Curran PF. Physiol Rev. 1970 Oct.50:637. [PubMed: 4919599]
3. Wright EM, Loo DD, Hirayama BA, Turk E. Physiology (Bethesda). 2004 Dec.19:370. [PubMed: 15546855]
4. Wright EM, Hirayama BA, Loo DF. J Intern Med. 2007 Jan.261:32. [PubMed: 17222166]
5. Reed-Tsur MD, De la Vieja A, Ginter CS, Carrasco N. Endocrinology. 2008 Jun.149:3077. [PubMed: 18339708]
6. Hediger MA, Coady MJ, Ikeda TS, Wright EM. Nature. 1987 Nov-Dec;330:379. [PubMed: 2446136]
7. Dohan O, et al. Endocr Rev. 2003 Feb.24:48. [PubMed: 12588808]
8. Jung H. FEBS Lett. 2002 Oct 2.529:73. [PubMed: 12354616]
9. Loo DD, Hirayama BA, Karakossian MH, Meinild AK, Wright EM. J Gen Physiol. 2006 Dec. 128:701. [PubMed: 17130520]
10. Eskandari S, et al. J Biol Chem. 1997 Oct 24.272:27230. [PubMed: 9341168]
11. Mackenzie B, Loo DD, Wright EM. J Membr Biol. 1998 Mar 15.162:101. [PubMed: 9538503]
12. Turk E, Wright EM. J Membr Biol. 1997 Sep 1.159:1. [PubMed: 9309206]
13. Turk E, et al. J Biol Chem. 2000 Aug 18.275:25711. [PubMed: 10835424]
14. See supporting material on Science online.
15. Eskandari S, Wright EM, Kreman M, Starace DM, Zampighi GA. Proc Natl Acad Sci U S A. 1998 Sep 15.95:11235. [PubMed: 9736719]
16. Fu D, et al. Science. 2000 Oct 20.290:481. [PubMed: 11039922]
17. Dutzler R, Campbell EB, Cadene M, Chait BT, MacKinnon R. Nature. 2002 Jan 17.415:287. [PubMed: 11796999]
18. Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E. Nature. 2005 Sep 8.437:215. [PubMed: 16041361]
19. Bowie JU. Nat Struct Mol Biol. 2006 Feb.13:94. [PubMed: 16462808]
20. Hunte C, et al. Nature. 2005 Jun 30.435:1197. [PubMed: 15988517]
21. Yernool D, Boudker O, Jin Y, Gouaux E. Nature. 2004 Oct 14.431:811. [PubMed: 15483603]
22. Hediger MA, Mendlein J, Lee HS, Wright EM. Biochim Biophys Acta. 1991 May 7.1064:360. [PubMed: 1903656]
23. Kinch LN, Grishin NV. Curr Opin Struct Biol. 2002 Jun.12:400. [PubMed: 12127461]
24. Sujatha MS, Balaji PV. Proteins. 2004 Apr 1.55:44. [PubMed: 14997539]
25. Abramson J, et al. Science. 2003 Aug 1.301:610. [PubMed: 12893935]
26. Zou JY, Flocco MM, Mowbray SL. J Mol Biol. 1993 Oct 20.233:739. [PubMed: 8240551]
27. Loo DD, et al. Proc Natl Acad Sci U S A. 1998 Jun 23.95:7789. [PubMed: 9636229]
28. Panayotova-Heiermann M, Loo DD, Lam JT, Wright EM. Biochemistry. 1998 Jul 21.37:10522. [PubMed: 9671524]
29. Hilger D, Bohm M, Hackmann A, Jung H. J Biol Chem. 2008 Feb 22.283:4921. [PubMed: 18156179]
30. De la Vieja A, Reed MD, Ginter CS, Carrasco N. J Biol Chem. 2007 Aug 31.282:25290. [PubMed: 17606623]
31. Pirch T, Landmeier S, Jung H. J Biol Chem. 2003 Oct 31.278:42942. [PubMed: 12923181]
32. Eswar N, et al. Curr Protoc Protein Sci. 2007 Nov.:9. **Chapter 2**, Unit 2. [PubMed: 18429317]

33. Veenstra M, Lanza S, Hirayama BA, Turk E, Wright EM. Biochemistry. 2004 Mar 30.43:3620. [PubMed: 15035632]

34. Turk E, Gasymov OK, Lanza S, Horwitz J, Wright EM. Biochemistry. 2006 Feb 7.45:1470. [PubMed: 16445289]

35. Hirayama BA, et al. Biochemistry. 2007 Nov 20.46:13391. [PubMed: 17960916]



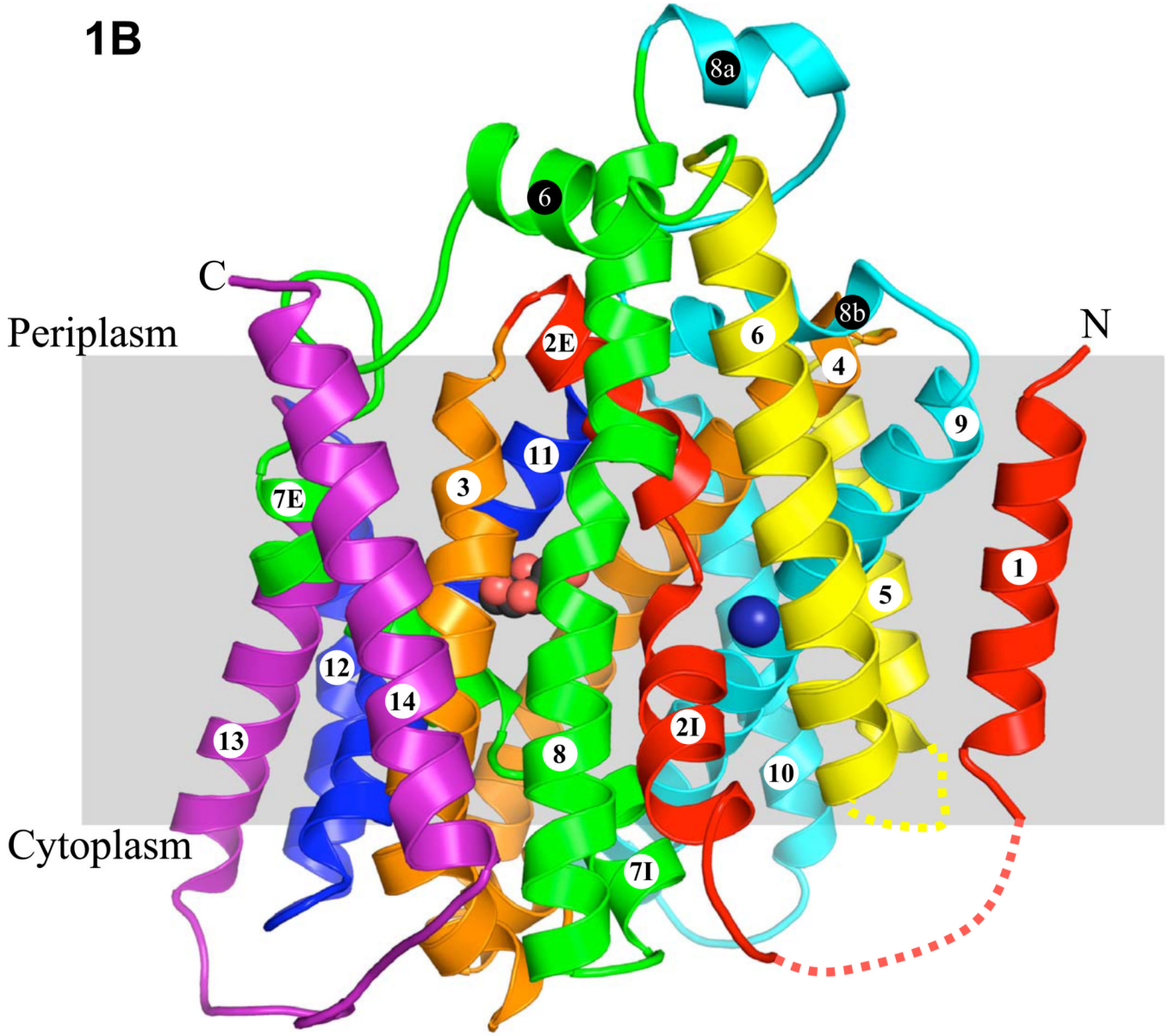
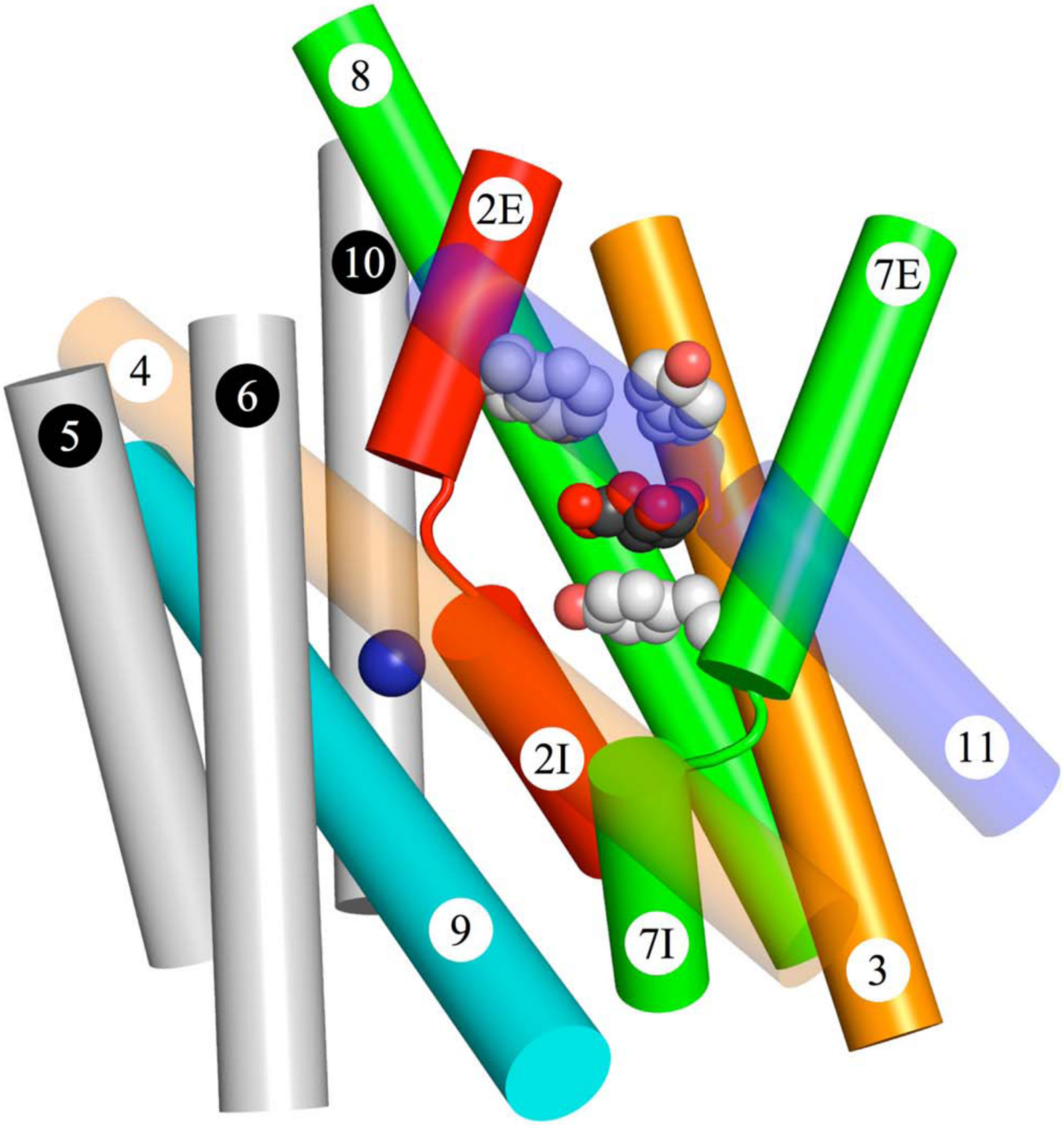
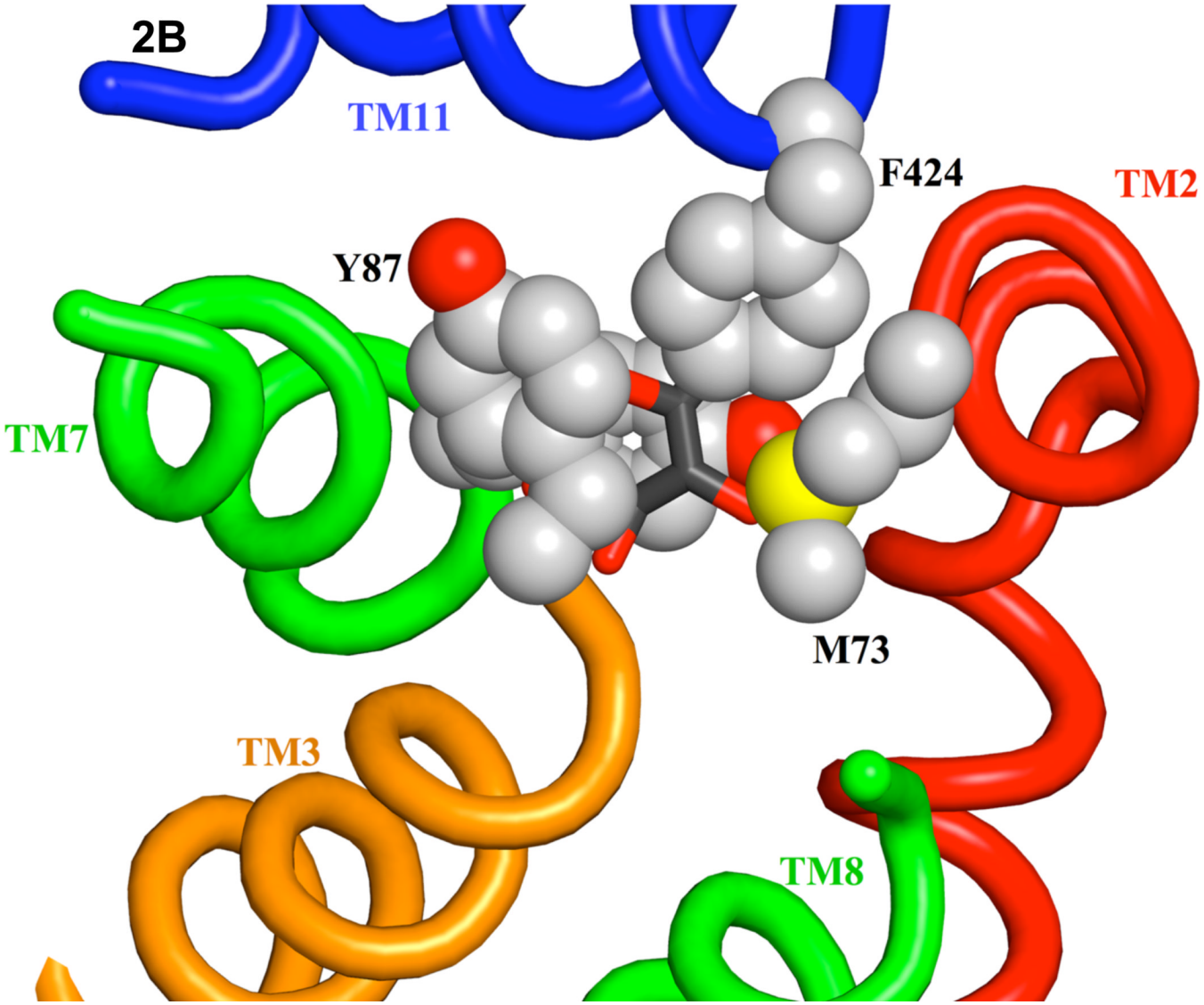


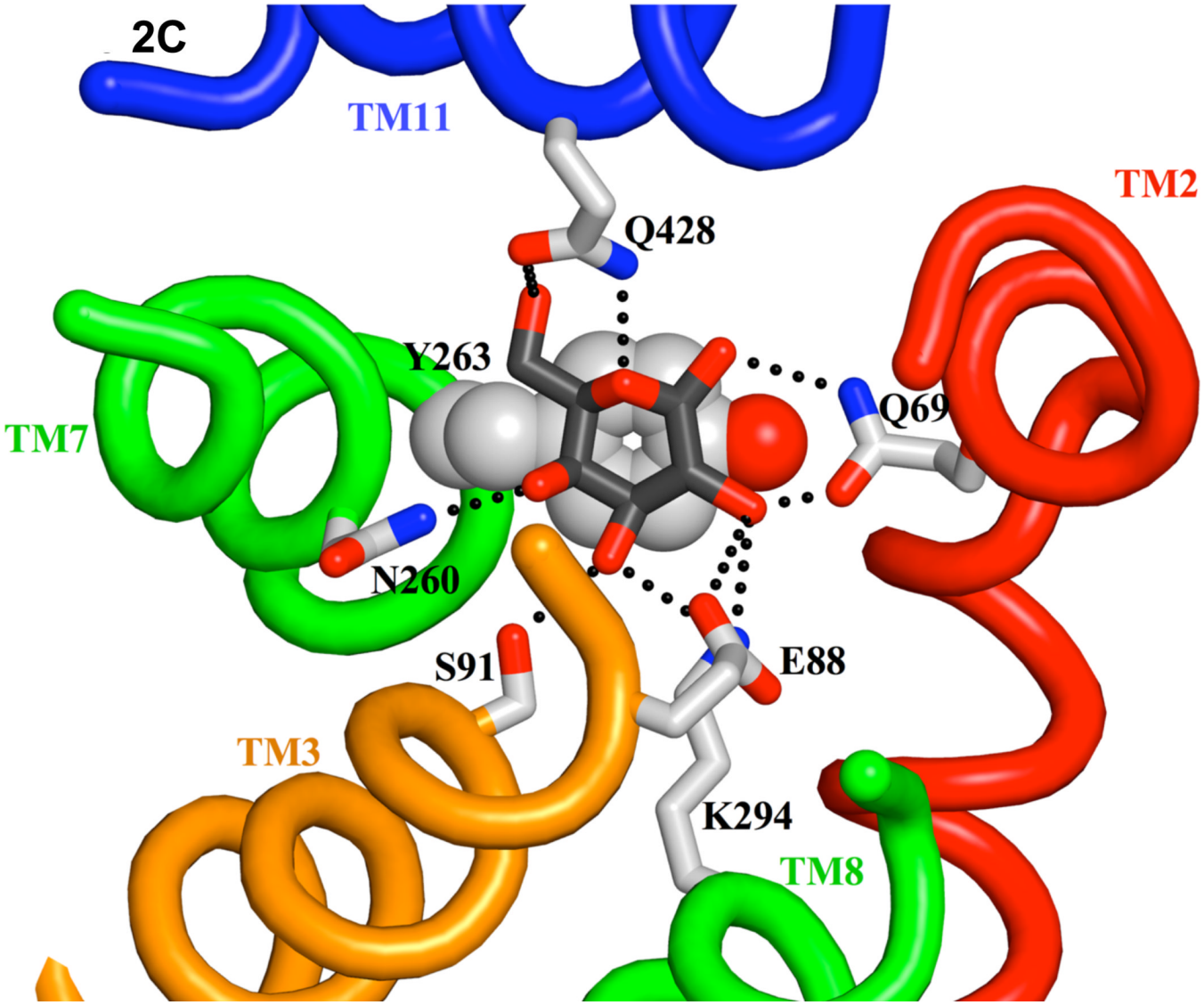
Fig. 1.

Structure of vSGLT. (A) Topology. The structure is colored as a rainbow from the N-terminus (red) to C-terminius (purple). The blue and red trapeziums represent the inverted topology of TM2-TM6 and TM7-TM11. The grey hexagon with red trim represents the galactose. Residues involved in sugar recognition, gate residues, and a proposed Na^+ site are shown in cyan, gray, and yellow circles. (B) Structure viewed in the membrane plane. The coloring scheme and numbering of α helices is the same as in Fig1A. Bound galactose is shown as black and red spheres for the C and O atoms. The proposed Na^+ ion is colored as a blue sphere. (C) Structure viewed from the intracellular side.

2A







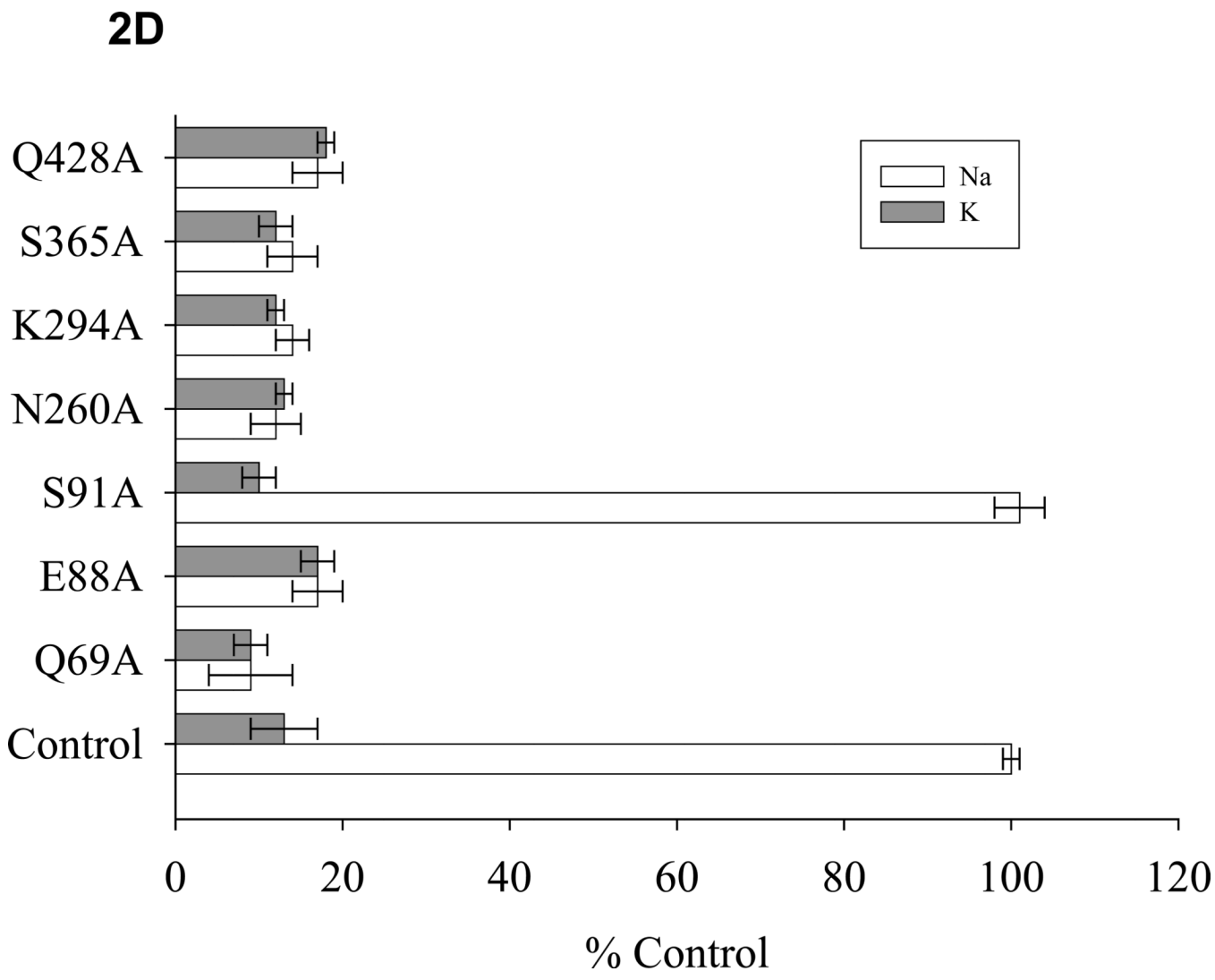
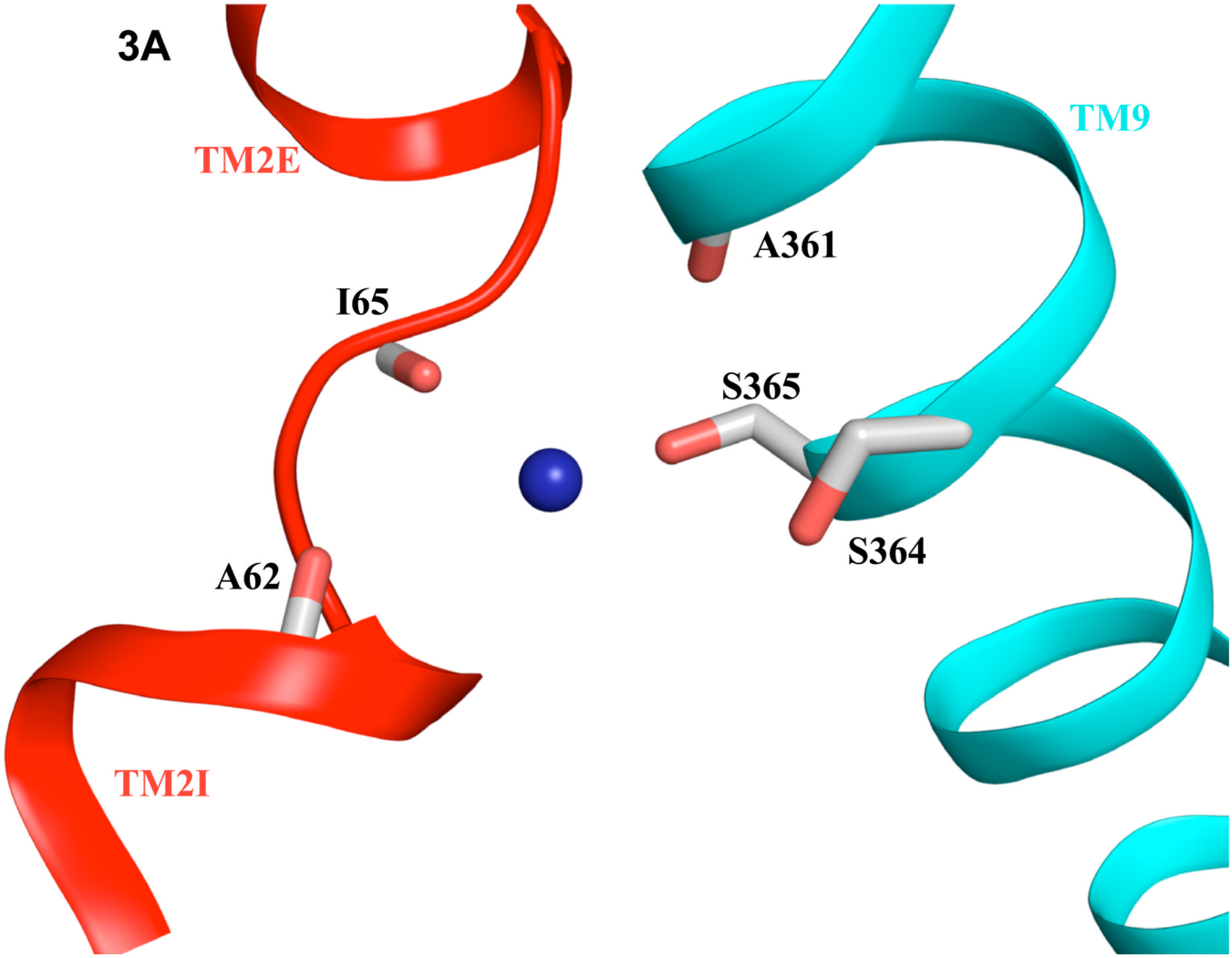


Fig. 2.

Galactose binding site. (A) Overview of the galactose and proposed Na^+-binding site viewed in the membrane plane maintaining the same color scheme as in Figure 1. (B) Hydrophobic gate residues (viewed from the extracellular side). The intracellular (Y263) and extracellular (M73, Y87 and F424) gates are shown as spheres, and the galactose is shown as sticks. (C) The galactose-binding site (same view as in B), with the extracellular hydrophobic gate residues removed to view the galactose-binding site. Residues in the galactose-binding site are displayed as sticks colored by atom type. Hydrogen bonds are depicted as black dashed lines. (D) D-galactose transport by vSGLT mutants in proteoliposomes. The results are expressed as % Control of that measured for p3C423 in 100 mM NaCl (~1.2 nmoles min^{-1} mg^{-1} protein). Standard error of the mean (SEM) is displayed for each experiment.



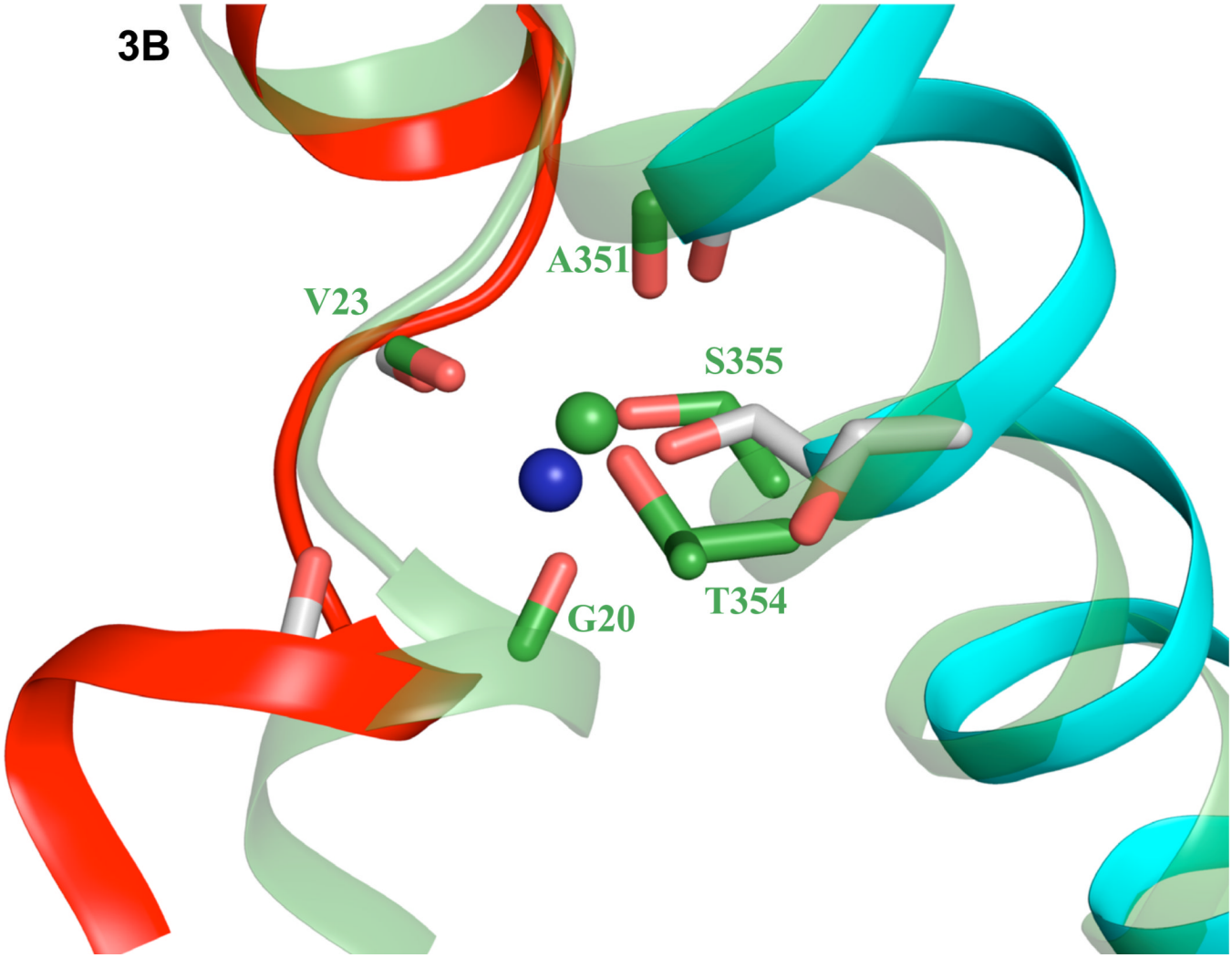


Fig. 3.

The proposed Na^+ binding site. (A) Residues in the Na^+-binding site are displayed as sticks colored by atom type, and the helices are colored as in Figure 1. (B) Superposition of the proposed Na^+ site of vSGLT on to the 2^{nd} Na^+ site from the LeuT structure (green). Alignment was performed using only helices 2, and 9 from vSGLT, and the corresponding helices from LeuT.

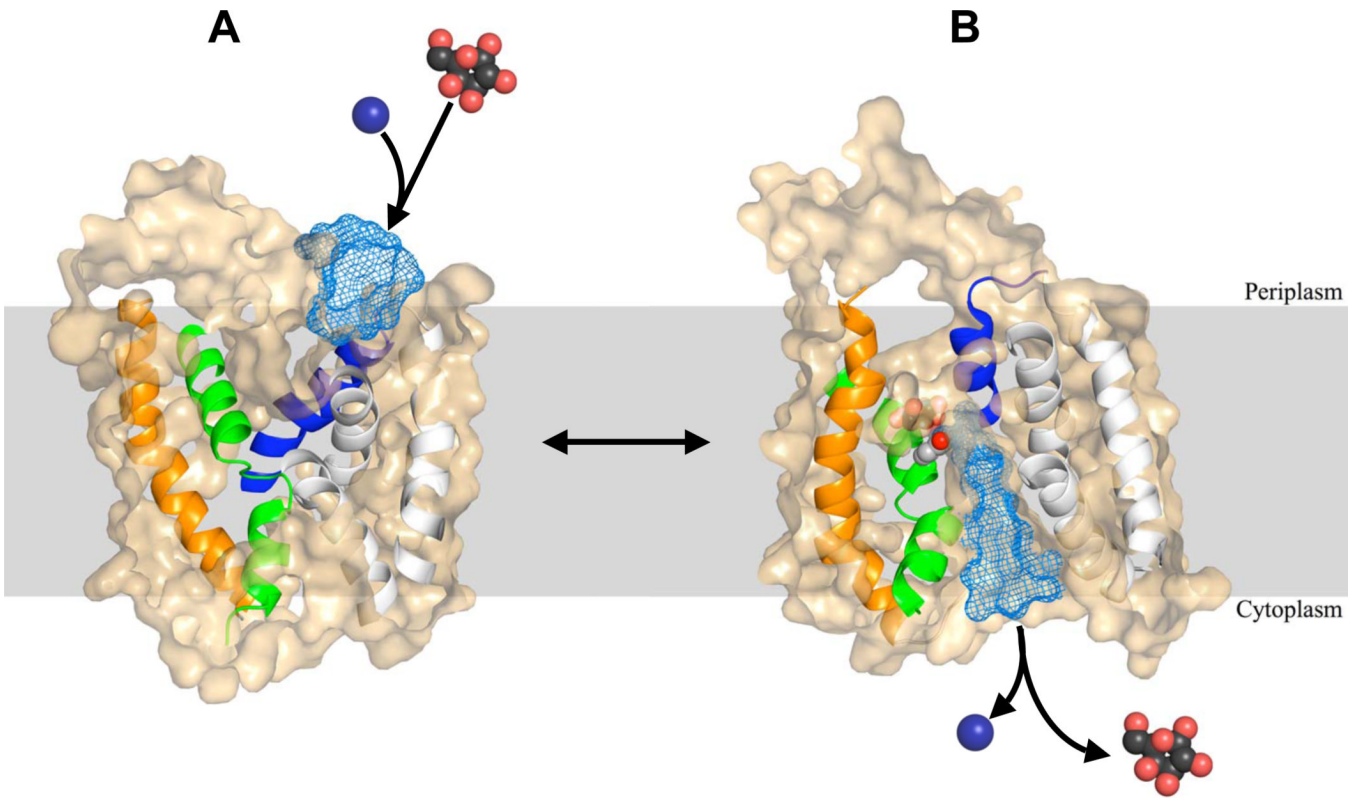


Fig. 4.

Alternating accessibility. (A) Slice through surface of the outward-facing model viewed from the membrane plane showing the extracellular cavity (blue mesh). (B) Slice through surface of the inward-facing structure of vSGLT in the membrane plane showing the intracellular cavity (blue mesh). Helices showing structural rearrangement are colored orange, green and blue for helices TM3, TM7 and TM11; helices with little movements are colored white. The surface is shown in beige. Galactose is shown as black and red spheres for the C and O atoms, respectively. Y263 and the Na^+ ion are colored as grey and blue sphere, respectively.