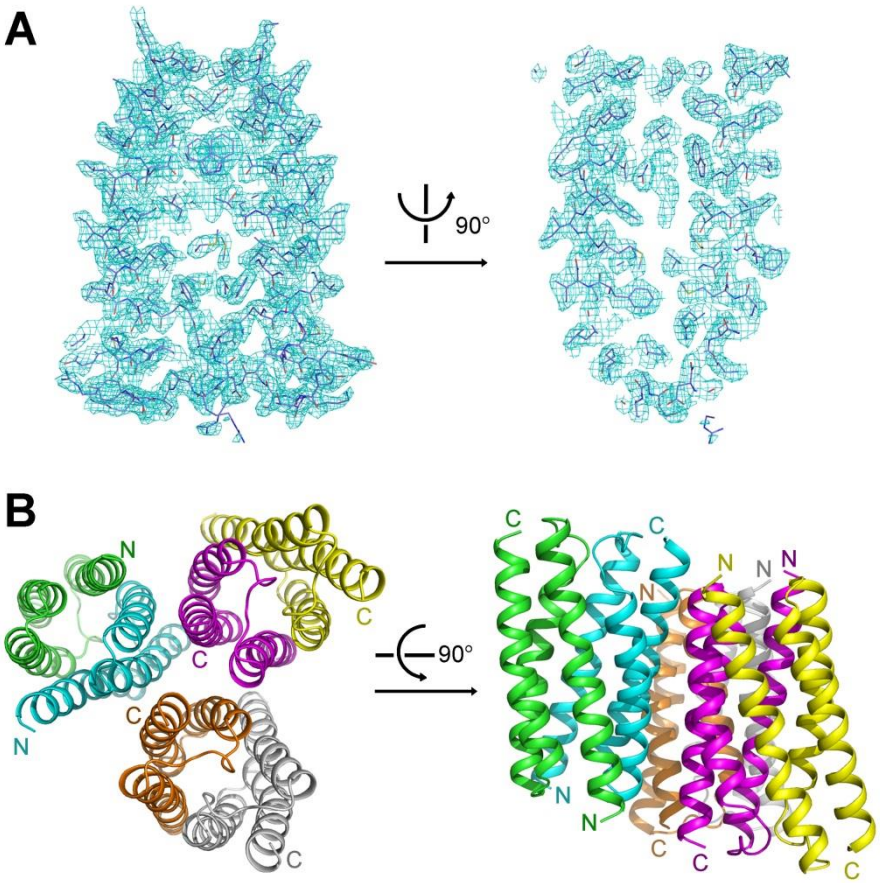
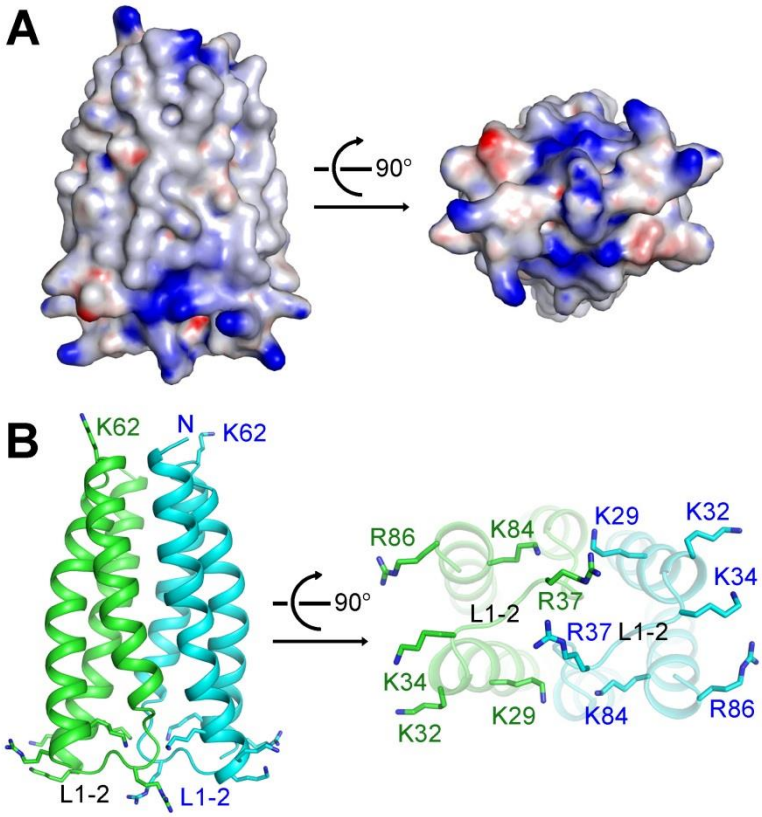


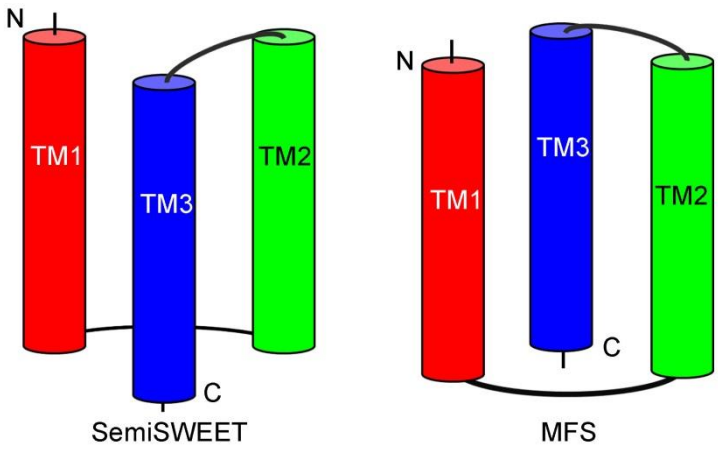
Supplementary information, Figure S1 Structure determination of TySemiSWEET.

(A) The refined structure of TySemiSWEET dimer is shown in the $2F_O$ - F_C electron density map. The map is contoured at 1.0 σ. **(B)** Each asymmetric unit has six molecules of TySemiSWEET that are organized into three dimers. One dimer exhibits an opposite orientation in contrast to the other two. Each color represents one molecule of TySemiSWEET. The amino termini and carboxyl termini are marked as “N” and “C”.



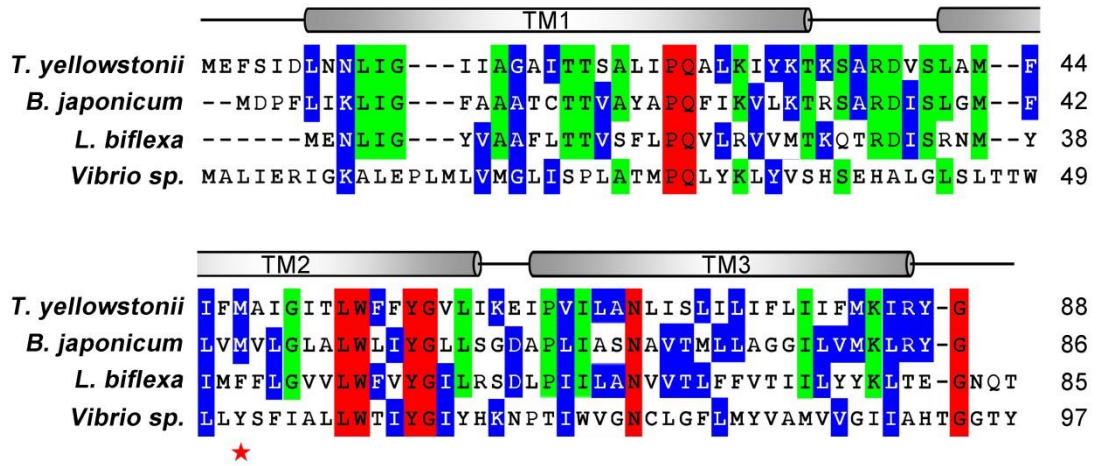
Supplementary information, Figure S2 Determination of the orientation of TySemiSWEET in membrane under the principle of the positive-inside rule [1].

(A) The electrostatic surface of TySemiSWEET generated by APBS [2]. Large positive charged areas are found at the bottom of TySemiSWEET dimer. **(B)** Distribution of positive residues in TySemiSWEET. Each TySemiSWEET protomer contains in total seven positively charged residues including five lysine residues and two arginine residues, which are shown as sticks. Notably, residues K29, K32, K34, K84, R37 and R86 are all located on or adjacent to L1-2, and only one residue, K62, is located on the opposite side. It indicates that L1-2 faces the cytoplasm.



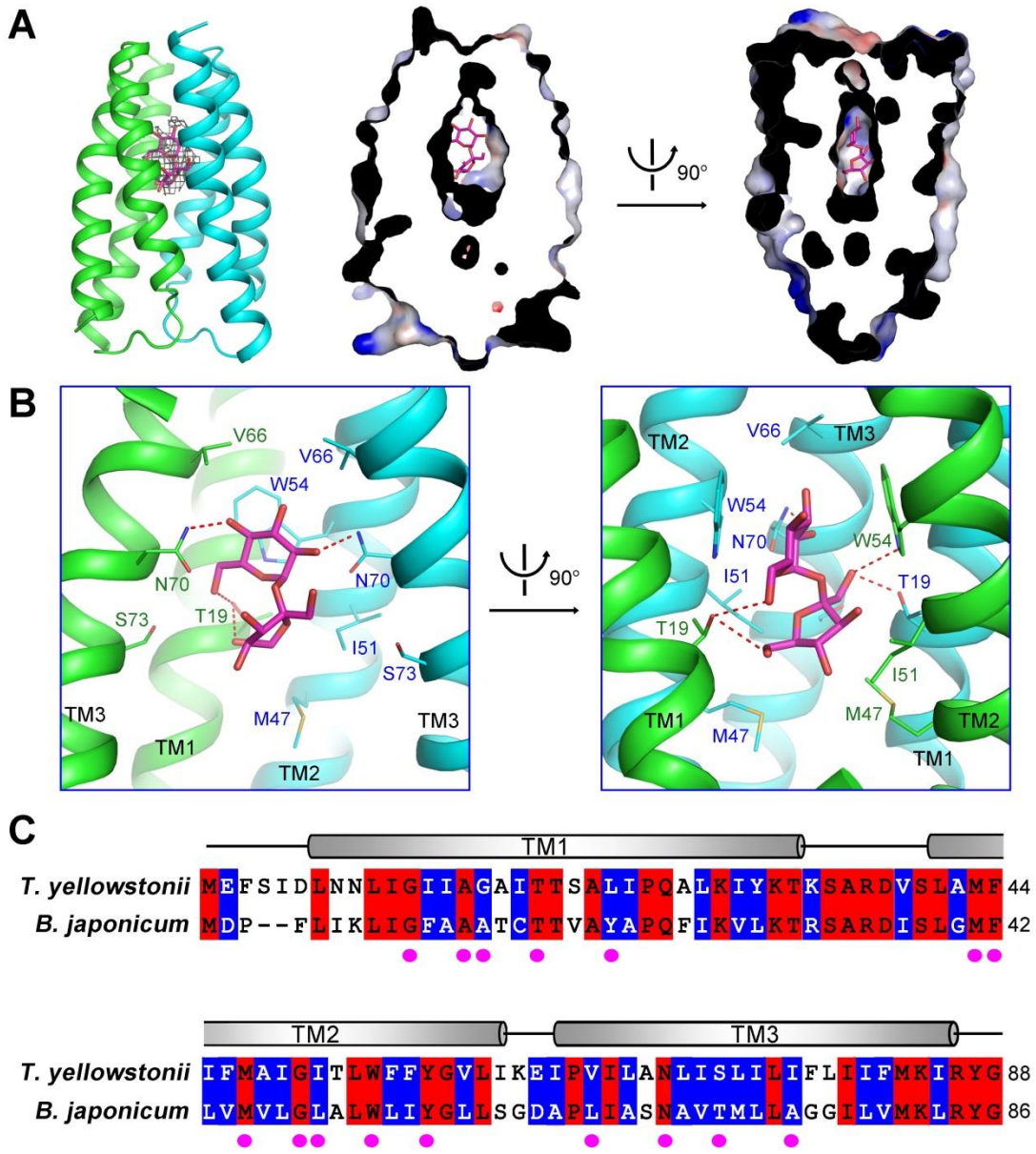
Supplementary information, Figure S3 Topology of the 3-helix bundles in SemiSWEET and MFS transporters.

SemiSWEET shows a similar but not identical topology compared to the 3-helix bundle in MFS. In SemiSWEET, TM3 locates in front of the plane defined by TM1 and TM2, whereas TM3 is located behind the TM1-2 plane in MFS.



Supplementary information, Figure S4 Sequence alignment of four SemiSWEETs from different species.

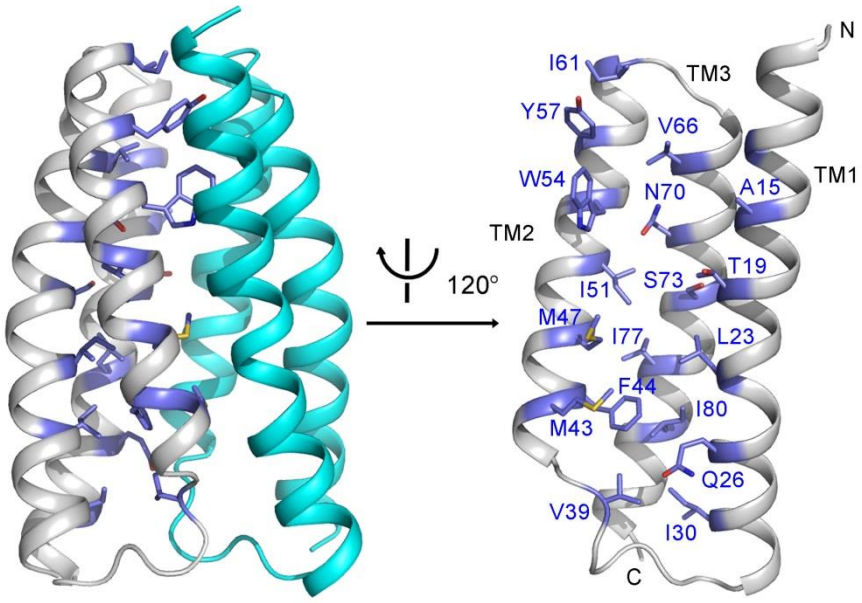
Red color indicates invariant residues in these SemiSWEETs; green and blue shades indicate the residues conserved in three and two of the four homologues, respectively. The key residue involved in substrate selectivity is indicated by red star at the bottom.



Supplementary information, Figure S5 The central cavity in TySemiSWEET can accommodate a saccharide molecule.

(A) Assignment of one sucrose molecule in the central pocket of TySemiSWEET. There is an obvious electron density in the central cavity of the TySemiSWEET dimer. The resolution and quality of the electron density does not allow accurate identification of the ligand. Given that the crystals were obtained in the presence of the putative substrate sucrose, we tentatively built into the electron density one sucrose molecule, which can be properly positioned in the

central cavity. **(B)** Residues in the central pocket that surround the sucrose molecule. In this scenario, the sucrose molecule interacts with T19, W54 and N70 through six hydrogen bonds. **(C)** Sequence alignment of TySemiSWEET and BjSemiSWEET. The invariant and conserved residues in the two proteins are shaded red and blue, respectively. The magenta circles indicate the residues that constitute the central pocket in TySemiSWEET.



Supplementary information, Figure S6 The transport pathway-facing residues in TySemiSWEET.

Residues in one protomer of TySemiSWEET that protrude to the potential transport path are shown in sticks.

References:

1 von Heijne G, Gavel Y. *Eur J Biochem* 1988; **174**: 671-678.

2 Baker NA, Sept D. Joseph S, *et al. Proc Natl Acad Sci U S A* 2001; **98**: 10037-10041.

Supplementary information, Table S1 Statistics of data collection and refinement.

Values in parentheses are for the highest resolution shell. $R_{merge}=\Sigma_h\Sigma_i|I_{h,i}-I_h|/\Sigma_h\Sigma_i I_{h,i}$, where I_h is the mean intensity of the *i* observations of symmetry related reflections of *h*.

$R=\Sigma|F_{obs}-F_{calc}|/\Sigma F_{obs}$, where F_{calc} is the calculated protein structure factor from the atomic model (R_{free} was calculated with 5% of the reflections selected).

Data	TySemiSWEET
Integration Package	HKL2000
Space Group	$P2_12_12_1$
Unit Cell (Å)	52.88, 96.16, 119.80
Unit Cell (°)	90, 90, 90
Wavelength (Å)	1.0000
Resolution (Å)	40~2.40 (2.49~2.40)
R_{merge} (%)	17.1 (54.9)
I/sigma	11.8 (2.3)
Completeness (%)	94.5 (95.3)
Number of measured reflections	136,846
Number of unique reflections	23,655
Redundancy	5.8 (5.3)
R_{work} / R_{free} (%)	25.09/29.36
No. atoms	
Protein	3,872
main chain	1,976
side chain	1,896
Water	-
Others	158
Average B value (Å^2)	
Protein	44.22
main chain	43.16
side chain	45.32
Water	-
Others	52.48
R.m.s. deviations	
Bonds (Å)	0.010
Angle (°)	1.443
Ramachandran plot statistics (%)	
Most favourable	97.1
Additionally allowed	2.9
Generously allowed	0.0
Disallowed	0.0

Supplementary information, Data S1

Methods

Protein expression and purification

The sequence of SemiSWEET from *Thermodesulfovibrio yellowstonii* (YP_002249334.1) was codon optimized specially for *E. coli* expression. The gene was cloned into pET21b (Novagen) with a hexahistidine (6xHis) tag directly linked to its carboxyl terminus. The plasmid was transformed into BL21. Cells were grown at 37 ℃ until to an optimal cell density at 600 nm (OD_{600}) of 1.0-1.2, and then the temperature was lowered to 18 ℃. Over expression of TySemiSWEET was induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 18 ℃ overnight. Cells were harvested by centrifugation, and resuspended in the lysis buffer containing 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, 20 mM sucrose and disrupted by sonication. After centrifugation at low-speed for 10 min, the supernatant was collected and centrifuged at 150,000 *g* for 1 h. Membrane fractions was resuspended in the lysis buffer, and then incubated with 2% (w/v) dodecyl-β-d-maltoside (DDM; Anatrace) at 4 ℃ for 1 hour and for an additional 20 min at 65 ℃. All the insoluble mass was removed by centrifugation at 150,000 *g* for 30 min. The supernatant was loaded to Ni^{2+}-nitrilotriacetate affinity column (Ni-NTA, Qiagen) at 4 ℃. The protein was eluted by the buffer containing 25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.04% DDM, 250 mM Imidazole and 20 mM sucrose. For the crystallization trials, the protein was further purified by gel filtration (Superdex-200 10/300 GL, GE Healthcare) with a running buffer containing 25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.04% DDM and 20 mM sucrose. The peak fractions were collected and concentrated to about 50 mg/mL for crystallization.

Crystallization

Purified TySemiSWEET protein was mixed with monoolein (protein:lipid ratio, 1:1.5) in syringe (Hamilton Company). Glass sandwich plates were used to set up trays using Gryphon LCP (Art Robbins). The crystals were grown to full size at 20 ℃ in two days in the buffer containing 2% (v/v) 3/4 EO/OH (Pentaerythritol ethoxylate), 100 mM Li_2SO_4 and 100 mM Sodium Citrate, pH 5.2. The crystals were harvested by mesh grid loops (MiTeGen), and then frozen in liquid nitrogen immediately for data collection.

Data collection and Structure determination

Native data of TySemiSWEET was collected at SPring-8 beamline BL32XU and processed with the package HKL-2000 [1]. Further processing was carried out with programs from the CCP4 suite [2]. Using atomic coordinates of the LbSemiSWEET protein (PDB code: 4QNC [3]) as the initial search model, the TySemiSWEET structure was determined by molecular replacement carried out by PHASER [4]. The atomic model was further rebuilt with COOT [5] and refined with PHENIX [6]. All of the figures were prepared using PyMOL [7].

References:

1 Otwinowski Z, Minor W. *Method Enzymol* 1997; **276**: 307-326.

2 Collaborative Computational Project. *Acta Crystallogr D Biol Crystallogr* 1994; **50**: 760-763.

3 Xu Y, Tao Y, Cheung LS, *et al. Nature* 2014.

4 McCoy AJ, Grosse-Kunstleve RW, Adams PD, *et al. J Appl Crystallogr* 2007; **40**: 658-674.

5 Emsley P, Cowtan K. *Acta Crystallogr D Biol Crystallogr* 2004; **60**: 2126-2132.

6 Adams PD, Grosse-Kunstleve RW, Hung LW, *et al. Acta Crystallogr D Biol Crystallogr* 2002; **58**: 1948-1954.

7 DeLano WL. *on World Wide Web http://www.pymol.org* 2002.