Supporting Information

Yaffe et al. 10.1073/pnas.1220497110

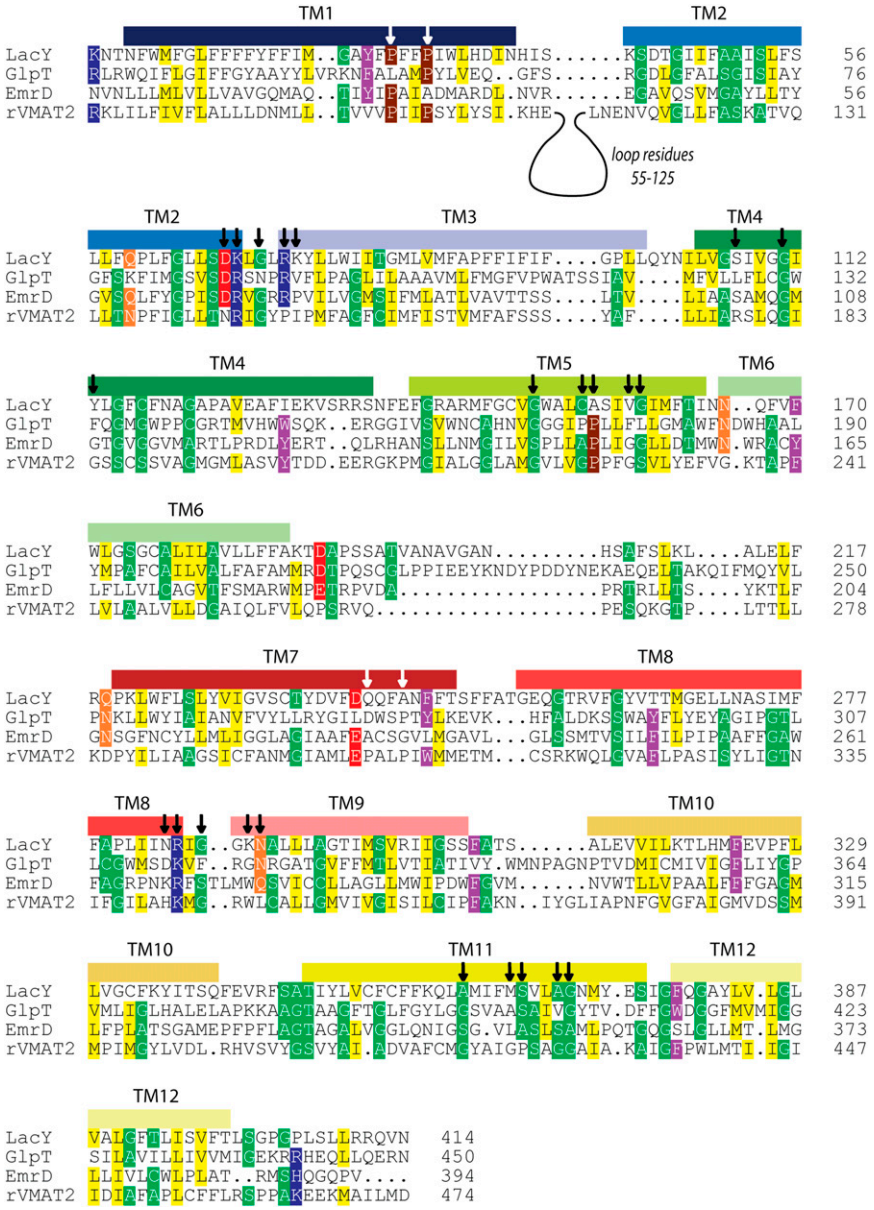

Fig. S1. Sequence alignment between rat vesicular monoamine transporter 2 (rVMAT2) and three major facilitator superfamily (MFS) transporters of known structure. Residues contributing to MFS or DHA12 motifs are marked with arrows. Conserved residues have been shaded according to the following color scheme: basic (H, R, K) in blue; acidic (D, E) in red; aromatic (F, Y, W) in purple; hydrophobic (L, I, V, M) in yellow; polar (N, Q) in orange; proline (P) in brown; and small (G, T, S, C, A) in green. The coloring of the transmembrane (TM) helices is as shown in Fig. 1.

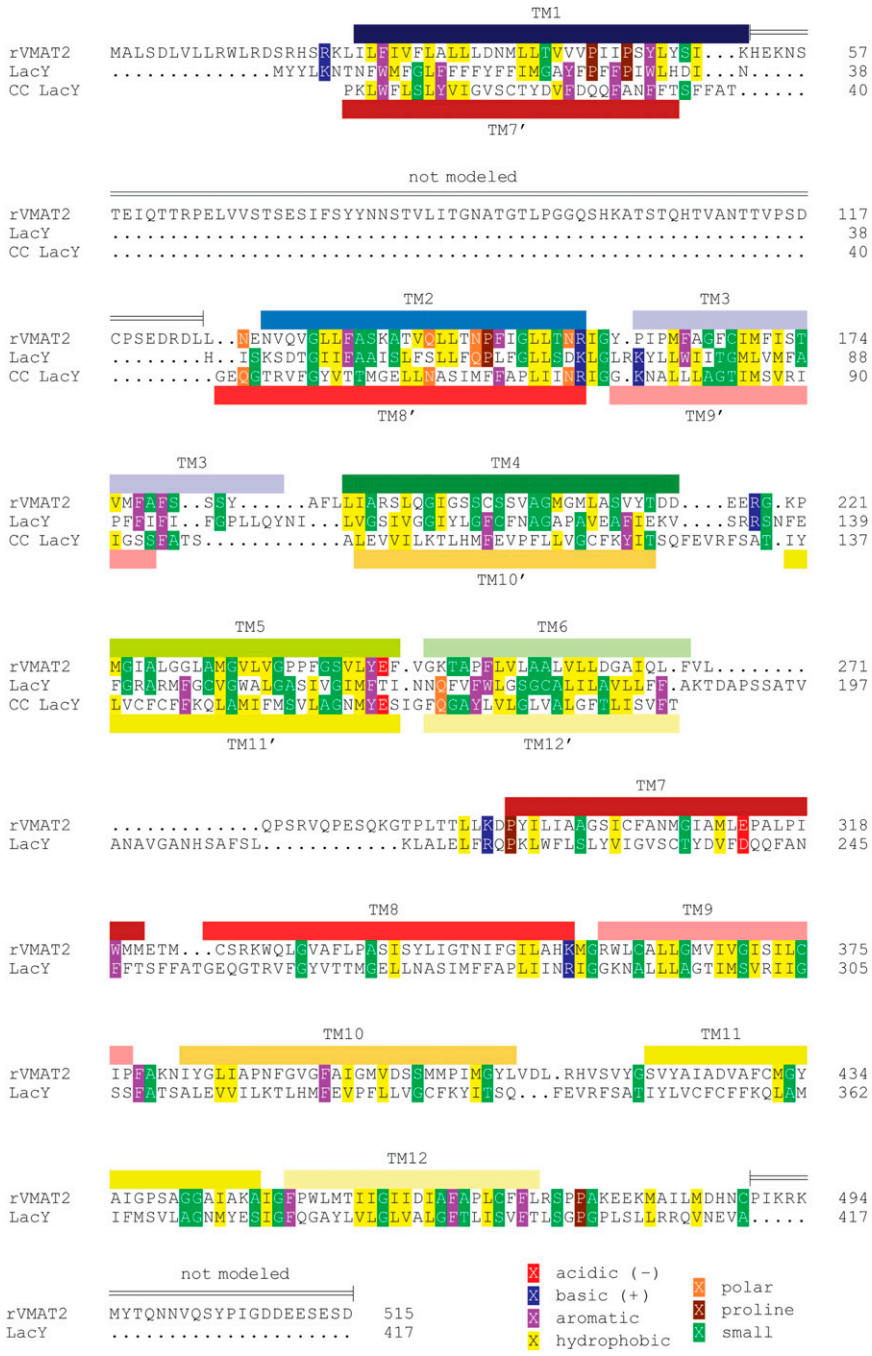

Fig. S2. Refined alignment between the lactose permease (LacY) and VMAT2 sequences used for modeling. The sequence labeled "CC LacY" consists of two copies of the C-terminal domain of LacY (see *Materials and Methods* in the main text). Residues contributing to DHA12 motifs are marked with arrows. See the legend of Fig. S1 for more details.

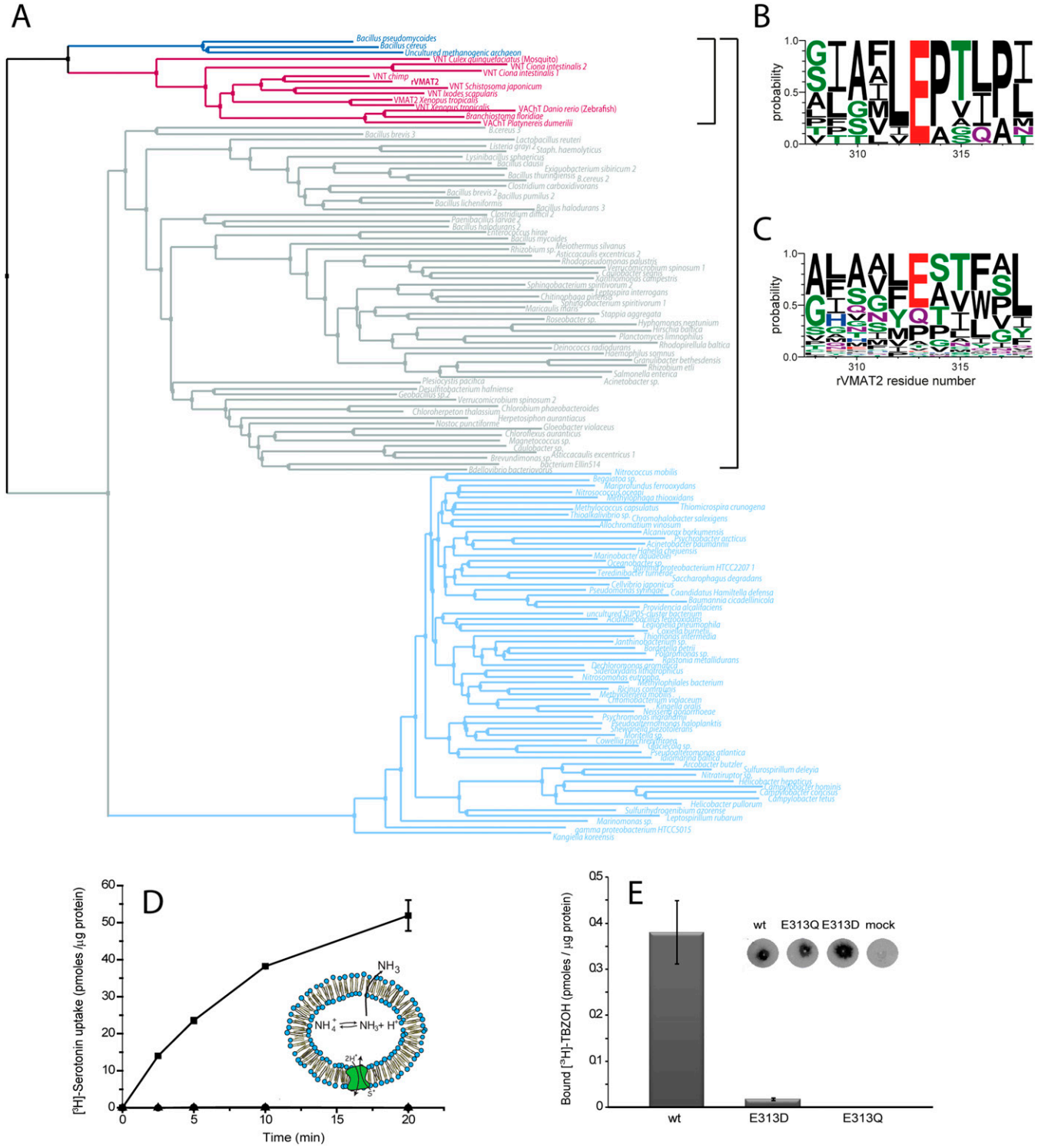

Fig. S3. Conservation and importance of E313. (*A*) Average-distance–based neighbor-joining phylogenetic tree for nonredundant sequence homologs of rVMAT2 identified using PSI-BLAST (see *Materials and Methods* in the main text). The sequences fall broadly into three clusters: VMATs (purple) and closely related prokaryotic MFS transporters (dark blue); more remote prokaryotic MFS transporters (gray); and distantly related prokaryotic MFS transporters (pale blue). Sequences were identified using PSI-BLAST (with five iterations and selecting the top 1,000 sequences, filtered by length, clustered to 60% sequence identity) and were multiply aligned using PRALINE (1). The tree was constructed from the multiple sequence alignment using the Phylip v3.67 programs *protdist* and *neighbor* (2) and visualized using Jalview v2.6 (3). (*B* and *C*) Sequence logos showing the probability of each amino acid type in the region containing E313 of rVMAT2, colored for acidic (red), basic (blue), small or polar (green), or other residue types (black). Logos were made by including only the branch containing eukaryotic VMATs (*B*) or by excluding the distantly related prokaryotic homologs (*C*). Logos were generated using Weblogo v3.3 (4). (*D* and *E*) Glu313 is irreplaceable. (*D*) [^{3}H]-Serotonin uptake into proteoliposomes. (*Inset*) Schematic representation of the proteoliposome system. The proton gradient was generated by loading the liposomes with NH_4^+ and diluting them into ammonium-free buffer. (*E*) [^{3}H]-dihydrotetrabenazine ([^{3}H]-TBZOH) binding to proteoliposomes. (*Inset*) Quantitation of protein amounts in proteoliposomes using a dot blot assay. Results presented are from duplicate experiments; errors bar indicate SD. Experiments were repeated at least twice.

1. Pirovano W, Feenstra KA, Heringa J (2008) PRALINETM: a strategy for improved multiple alignment of transmembrane proteins. *Bioinformatics* 24:492–497.
2. Felsenstein J (1989) PHYLIP–Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164–166.
3. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25:1189–1191.
4. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: A sequence logo generator. *Genome Research* 14:1188–1190.

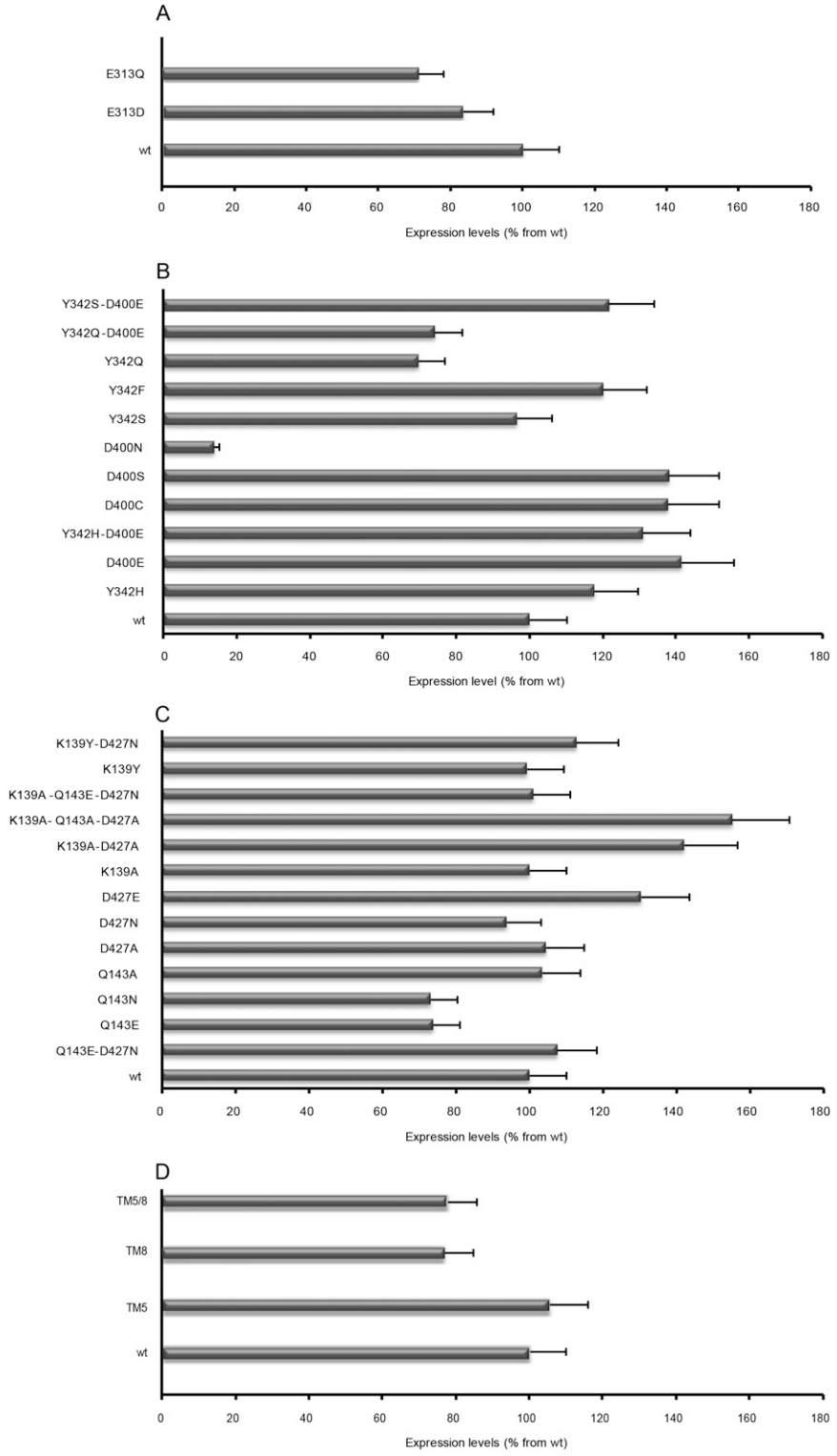

Fig. S4. Expression levels of rVMAT2 mutants. Membranes were analyzed, and expression levels quantified as described in *Materials and Methods* in the main text. Three mutants showed no transport activity but expressed to very low levels and therefore were not characterized further: D400N (shown in the figure) and E313A and Q143N/D427E (expression too low to quantitate). (*A–D*) The results of four independent experiments are shown. In each experiment, wild type cells were used as a control and set at 100% value.

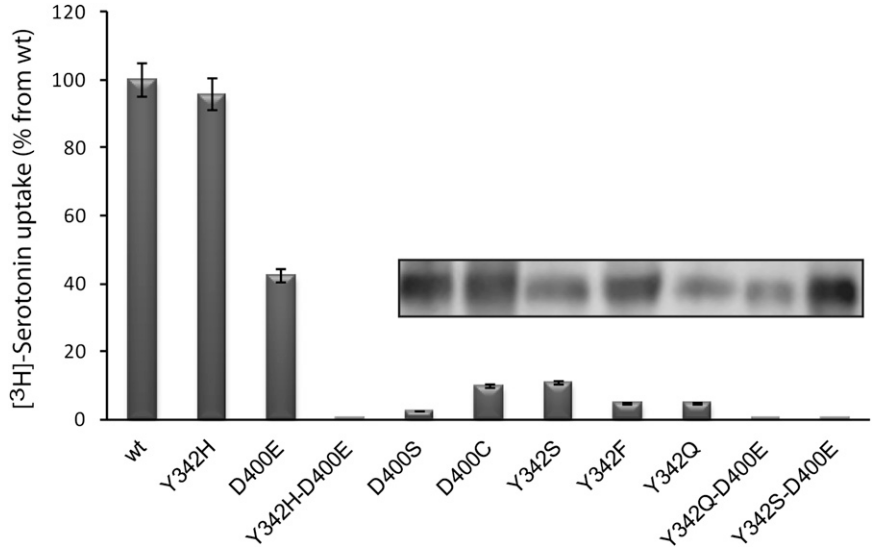

Fig. S5. [^{3}H]-Serotonin uptake into HEK293 whole cells. (*A*) [^{3}H]-Serotonin uptake as described in Fig. 3. (*Inset*) Western blot. Membranes were analyzed as described in *Materials and Methods* in the main text. Results presented are from duplicate experiments; error bars indicate SE. Experiments were repeated at least twice.

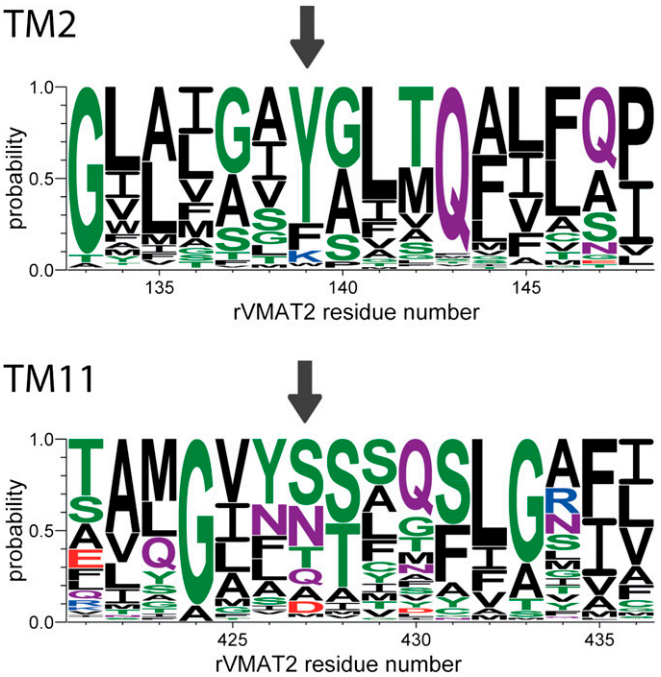

Fig. S6. Sequence conservation of positions in the polar cluster. Sequence logos quantifying the probability of each amino acid type in the region containing K139 (TM2) and D427 (TM11) of rVMAT2. Key positions are indicated by arrows. Sequences were taken from the homologs of VMAT2 shown in Fig. S3, excluding the most distantly related prokaryotes. Residues are colored as in Fig. S3*B*.

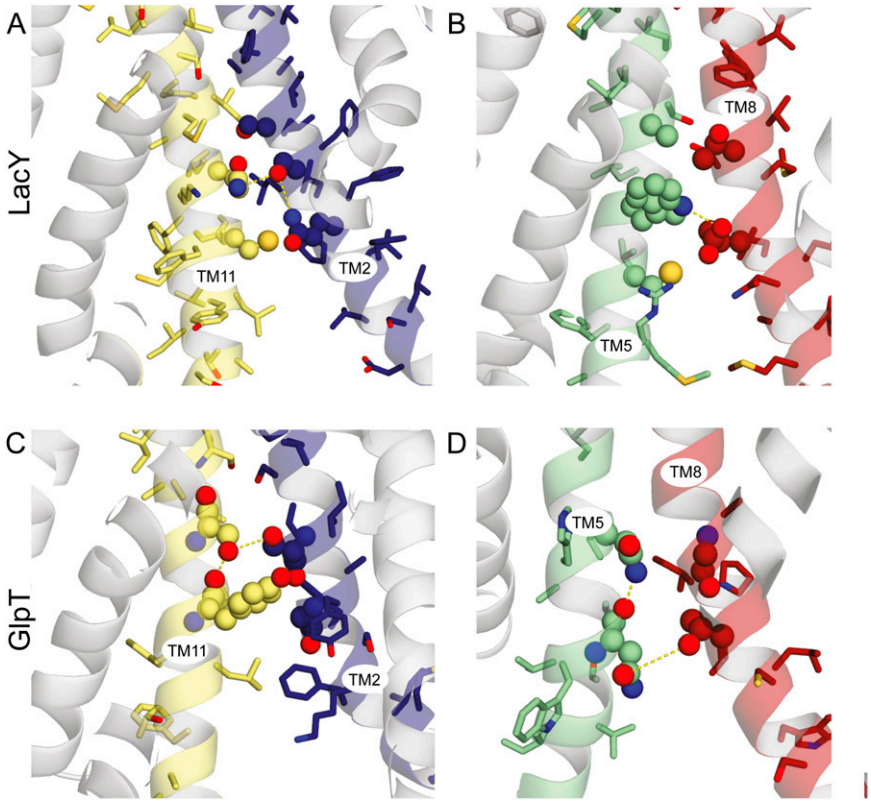

Fig. S7. Nature of the interaction networks between TM2/11 and TM5/8 in other MFS members. Crystal structures of LacY [Protein Data Bank (PDB) ID code 1PV7] (*A* and *B*) and GlpT (PDB ID code1PW4) (*C* and *D*) in the cytoplasm-facing state are viewed along the plane of the membrane from inside the central cavity and oriented with the cytoplasm toward the bottom of the page. Residues in the relevant TMs are shown as sticks. Residues at the putative hinge points (Table S1) are shown as spheres. (*A*) Residues S53, S56, Q60, C355, and Q359 from TM2/11 in LacY. (*B*) Residues C148, W151, A155, T265, and E269 from TM5/8 in LacY. (*C*) Residues S73, G77, Y393 and S397 from TM2/11 in GlpT.(*D*) Residues N162, N166, G302, and T306 from TM5/8 in GlpT.

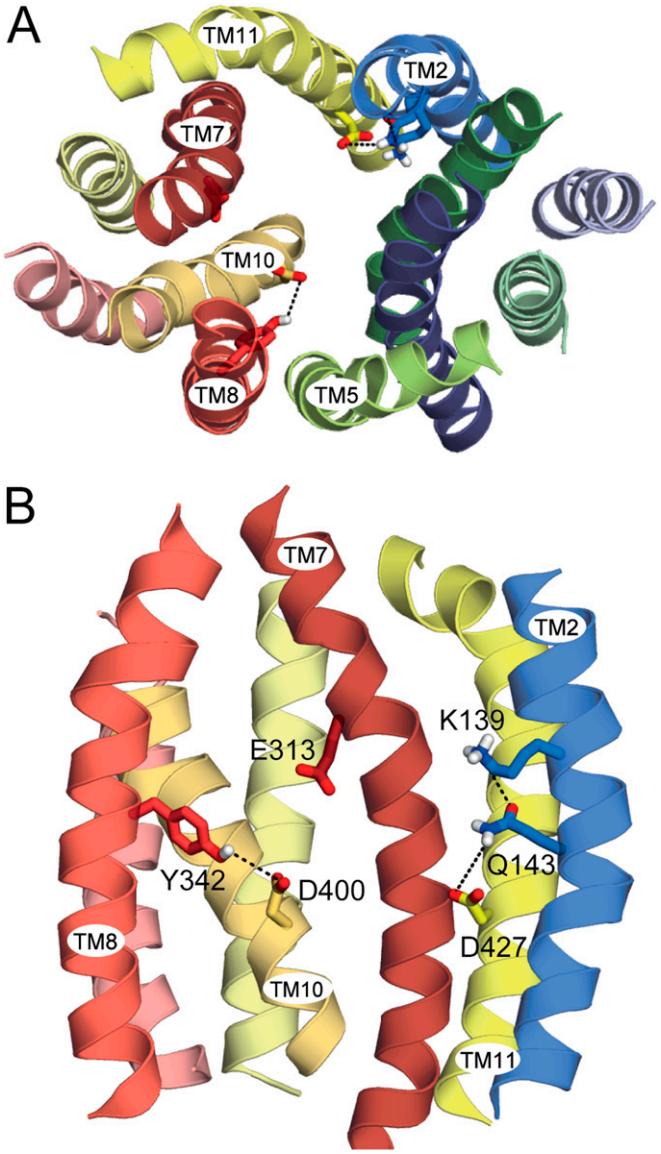

Fig. S8. Model structure of VMAT2 in an outward-open conformation. The protein is viewed from the cytoplasm (*A*) and along the plane of the membrane (*B*) with the cytoplasm toward the bottom of the page. Helices are shown as cartoons, colored according to Fig. 1, with the specific residues studied here shown as sticks. Predicted interactions between side chains are indicated by dashed lines.

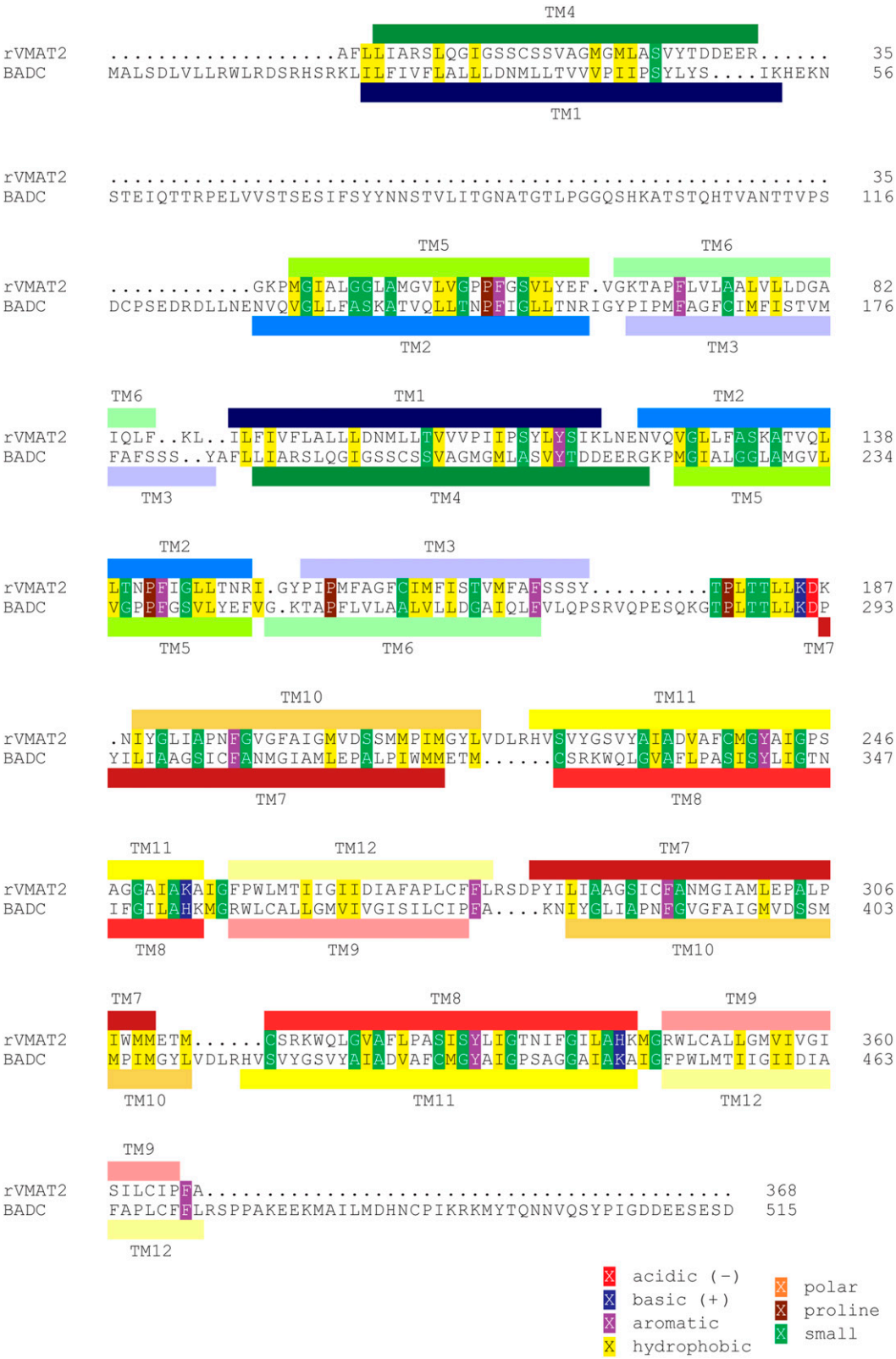

Fig. S9. Refined alignment between template and model sequences for the vesicle lumen-facing repeat-swapped model of rVMAT2. The sequence labeled "BADC" consists of the rVMAT2 repeat units in the order B, A, D, then C. Conserved residues have been shaded as in Fig. S1.

Table S1. Conservation of residues at the interface between the N- and C-terminal domains in different MFS subfamilies

Protein	TM2	Score*	R	TM11	Score	R	TM5	Score	R	TM8	Score	R
VMAT2	K139	−1.07	9	D427	−1.08	9	V233	−0.49	7	F335	−0.85	8
	Q143	−0.98	9				L234	−0.73	8	L336	−0.58	7
XylE	I69	−0.54	7	W416	−0.97	9	Q168	−1.03	9	G322	−1.01	9
	C68	−0.66	8	N419	−0.94	9	F169	−0.42	7	N325	−0.90	9
	L65	−0.69	8	Y420	0.37	4	I172	−0.91	9	T329	−0.92	9
$PepT_{St}$	A64	−0.82	8	S431	−0.41	7	N156	−1.09	9	P329	−0.92	9
	S65	−0.64	8	S432	−0.68	8	A159	−0.43	7	I332	−0.85	8
	Y68	−0.58	7	S435	−0.79	8				M333	−0.64	8
LacY	S53	−0.71	8	C355	−0.32	6	C148	−0.97	9	T265	−0.70	8
	S56	−0.71	8	Q359	−0.77	8	W151	−0.51	7	E269	−0.54	7
	Q60	−1.10	9				A155	−0.70	8			
GlpT	S73	−0.52	7	Y393	−1.13	9	N162	−1.10	9	G302	−0.90	9
	G77	−1.02	9	S397	−0.58	7	N166	−1.06	9	T306	−0.86	8
FucP	Y74	−1.15	9	V388	−0.70	8	T160	−0.77	8	F309	−0.41	7
	P78	−0.95	9	I391	−1.02	9	S163	−0.96	9	R312	−1.15	9

R, normalized ranking; rankings range from 0 to 9. For reference, the ConSurf score for E313 of rVMAT2 is −0.87, with a ranking of 8. ConSurf v3.0 was used with default parameters for collecting sequence homologs. For further analysis, two or three positions were selected for each helix approximately at the midpoints where the two domains would be expected to be in contact in both conformations.

*ConSurf scores were calculated for MFS representatives for which structures are known. PDB ID codes are 4GBY (XylE), 4APS ($PepT_{St}$), 1PV7 (LacY), 1PW4 (GlpT), and 3O7P (FucP).