Supporting Information

Crystal structure of Ca^{2+}/H^{+} antiporter protein YfkE reveals the mechanisms of Ca^{2+} efflux and its pH regulation

Mousheng Wu[1], Shuilong Tong[1], Sandro Waltersperger[2], Kay Diederichs[3], Meitian Wang[2] & Lei Zheng[1]

[1]Center for Membrane Biology, Department of Biochemistry and Molecular Biology, the University of Texas Houston Medical School, Houston, TX 77030, USA. [2]Swiss Light Source, Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland. [3]Department of Biology, University of Konstanz, D-78457 Konstanz, Germany

SI Appendix, Methods and Materials

Protein preparation *YfkE* was cloned into the pET22b vector from *Bacillus subtilis* genome. A hexa-His tag was inserted at the residue 188 within the cytoplasmic loop between TMs 5 and 6 to facilitate protein purification and the generated protein is referred to as YfkE wild type. To improve protein expression and crystallization, the mutation K116A was introduced within the cytoplasmic loop between TMs 3 and 4. The native YfkE protein was expressed in *E.coli* strain BL21 (DE3) in autoinduction medium (1) at 37 °C for 3 h followed by overnight incubation at 25 °C. Cells were harvested and homogenized in phosphate buffer containing 20 mM sodium phosphate, pH 7.4, 500 mM NaCl and 20 mM imidazole. Cell rupture was carried out by three passages through a C3 homogenizer (Avestin) at 15,000 p.s.i. Cell debris was removed by centrifugation. The supernatant was pelleted by ultracentrifugation for 1 h. Membrane fractions suspended in the same phosphate buffer as described above were solubilized by addition of 1% (*w/v*) n-dodecyl-β-maltoside (DDM, Anatrace) at 4 °C. After 1 h incubation, the solubilized solution was centrifuged at 40,000 rpm for 25 min. The supernatant was loaded on to a Ni-NTA affinity resin (GE healthcare). The resin was washed with phosphate buffer containing 60 mM imidazole and 0.05% DDM (*w/v*). The YfkE protein was eluted with phosphate buffer containing 400 mM imidazole and 0.05% DDM. The eluted protein was further purified by SEC using a Superdex-200 10/300 GL column (GE Healthcare) equilibrated in buffer containing 20 mM sodium phosphate, 500 mM NaCl, 0.05% DDM (*w/v*), pH 7.4. The final protein solution was concentrated to 8 mg mL^{-1} and stored at -80°C before crystallization. An additional mutation L77M was introduced to crystallize Se-Met protein. Se-Met protein of YfkE was expressed in the BL21(DE3) C43 *E.coli* strain using a published protocol (2), and purified using similar

approach to that for the native protein. After protein eluted from Ni-NTA resin, 5mM DTT was added to the eluted protein solution and maintained in all subsequent buffers.

Protein crystallization Both native and selenomethionine-substituted YfkE proteins were crystallized in buffer containing 23-25% PEG 400, 200 mM ammonium sulfate, 20mM NaCl, 0.1M sodium acetate, pH 4.0 and 3% hexanediol using the sitting-drop vapor diffusion method at room temperature. 2 mM Ca^{2+} was supplemented to the native protein crystallization buffer. Both crystals diffract to a maximum 8 Å resolution, estimated by in house Rigaku X-ray source at room temperature. The crystals were further treated using the crystal dehydration and salvage kit (Jena Biosciences) for 30 min at room temperature. This dehydration treatment is critical to improve the crystal quality to 4 Å resolution. The dehydrated crystals were cryo-protected in Paraffin oil and flash frozen quickly under liquid Nitrogen.

Data collection, data processing and structure determination The native YfkE dataset was collected at the Swiss Light Source (SLS, Switzerland) and the MAD datasets were collected at the Advanced Light Source beamline 4.2.2 (Berkeley, CA, USA). Both the native and the Se-Met crystals diffract to the highest resolution of 3.1 Å. Data were processed using the program *XDS* (3). Statistics of data collection and data processing are given in *SI Appendix*, Table S1. The data were initially processed to the space group of $R3_2$ with one molecule in the asymmetric unit. Both exhibit nearly perfect twining (~45%), estimated by the *phenix.xtriage* program (4). The initial phase was calculated using the MAD method and density modification by the *AutoSHARP* program (5). Each YfkE peptide has 12 native methionine residues. Ten strong Se sites were unambiguously localized. A partial YfkE model containing the main chains of 11 α helices was built using *Coot* (6) and refined in the $R3_2$ space group using *Refmac* (7) till R / R_{free} stabilized at 38% / 42%. The electron density map was gradually improved after several rounds of the MR-SAD (Molecular Replacement combined with SAD phasing) approach using *Phaser* (8) and model building with the peak data. To help model building and to verify the YfkE model, five additional methionine mutants on different TM helices were also crystallized. Their peak datasets were collected at ALS or SLS (*SI Appendix,* Table S2). New heavy atom sites were localized by calculating their anomalous difference maps using *AutoSHARP* (5) in the $R3_2$ space group. The native YfkE structure was determined by the molecular replacement method in the *R3* space group using *Phaser* (8). Twin refinement was performed in the *R3* space group by

Refmac (7) with twin law (*k, h, -l*). The final models were refined to R / R_{free} values of 22% / 26% for the native structure and 24% / 27% for the Se-Met L77M structure. All residues fall into the allowable region of Ramachandran plot. Statistics of structure refinements are given in *SI Appendix* Table S1. The structures of YfkE native and Se-Met proteins have been deposited to Protein Data Bank (PDB accession codes: 4KJR & 4KJS). All structural figures were prepared using *PyMOL* (9).

Disulfide crosslinking Disulfide crosslinking experiments were performed using an iodine-catalyzed approach described by Kaback for LacY (10). *E.coli* strain BL21(DE3) cells expressing YfkE wild type or cysteine mutants were suspended in lysis buffer (20 mM sodium phosphate, 500 mM NaCl, 10μg mL-1 lysozyme, pH7.4) and then ruptured by sonication. Crosslinking was triggered by adding aqueous iodine (final concentration of 0.5 mM) into the whole cell lysate at room temperature. After 30 min incubation, reactions were stopped by adding 10 mM N-ethylmaleimide and then were separated by SDS-PAGE. Crosslinking was visualized by western blot using anti-peta-His antibody horseradish peroxidase-conjugate.

Ca^{2+} transport assay Ca^{2+} efflux assay was performed by measuring Ca^{2+} uptake into everted membrane vesicles. YfkE everted vesicles were prepared using a method described by Rosen and Tsuchiya with modification (11). Briefly, *E.coli* strain BL21(DE3) cells carrying the YfkE expression vector were grown in Luria broth medium at 37°C. After cell density reached 0.4 at OD600, protein expression was induced by adding 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 2 h at 25°C. Cells were then harvested and washed with TKDS buffer (10 mM Tris-HCl, 140 mM KCl, 0.5 mM DTT, 250 mM sucrose, pH 7.3). Everted vesicles were generated by single passage of cells through a C3 homogenizer (Avestin) at 4,000 p.s.i. After the cell debris was removed, the supernatants were centrifuged to pellet the membrane fractions. The vesicles were homogenized in TKDS buffer and quick frozen in liquid Nitrogen for use.

Thawed vesicles were diluted to 0.13 mg mL^{-1} total protein concentration with TKDS buffer (pH 8.0) containing 5 mM potassium phosphate. Prior to assays, the vesicles were incubated with 5 mM NADH for 10 min. Reaction was triggered by addition of $^{45}CaCl_2$ into the vesicles. Samples were taken at various times, filtered through a nitrocellulose membrane (0.22

μm) on a Millipore filtration manifold and immediately washed with 10 ml of TKDS buffer, pH 7.3. The filters were air-dried and counted in a liquid scintillation counter. *i*) For time-dependent transport assay (*SI Appendix*, Fig. S2*A*), 500 μM substrate was used. *ii*) For cation inhibition assays (*SI Appendix*, Fig. S3), the vesicles were incubated with 200 μM (10 fold) or 2 mM (100 fold) of $CaCl_2$, LiCl, NaCl, $MgCl_2$, $MnCl_2$, $CdCl_2$, $CoCl_2$, $NiCl_2$ or $SrCl_2$, before 20 μM $^{45}Ca^{2+}$ was added to the reactions for 10 min. *iii*) For pH dependent transport measurement (*SI Appendix*, Fig. S2*B*), YfkE vesicles were diluted into MES-Tris buffer solutions at appropriate pHs prior to measurements. *iv*) For Michaelis Menten kinetic analyses (Fig. 4*C* and *SI Appendix*, Fig. S2*B*), the transport activity was measured at seven different [Ca^{2+}] ranged from 5 to 500 μM. Apparent K_m and V_{max} values were calculated using *Graphpad Prism*^{TM} software. All assays were performed in triplicate. Everted vesicles prepared with the cells hosting empty vector were used as control.

Circular dichroism (CD) spectroscopy Prior to CD spectroscopic analysis, Ni-NTA purified YfkE wild type or mutant proteins were passed through a P-10 desalting column (GE healthcare) equilibrated with phosphate buffer (20 mM sodium phosphate, 400 mM Na_2SO_4, 0.05% DDM, pH 7.4) to remove Cl^-. The protein concentration of all samples was adjusted to 0.2 mg mL^{-1}. CD spectra were collected at room temperature over a wavelength range from 190 to 260 nm with a Jasco J-815 CD spectrometer using a 0.02-cm cylindrical cell.

SI Appendix, Figures

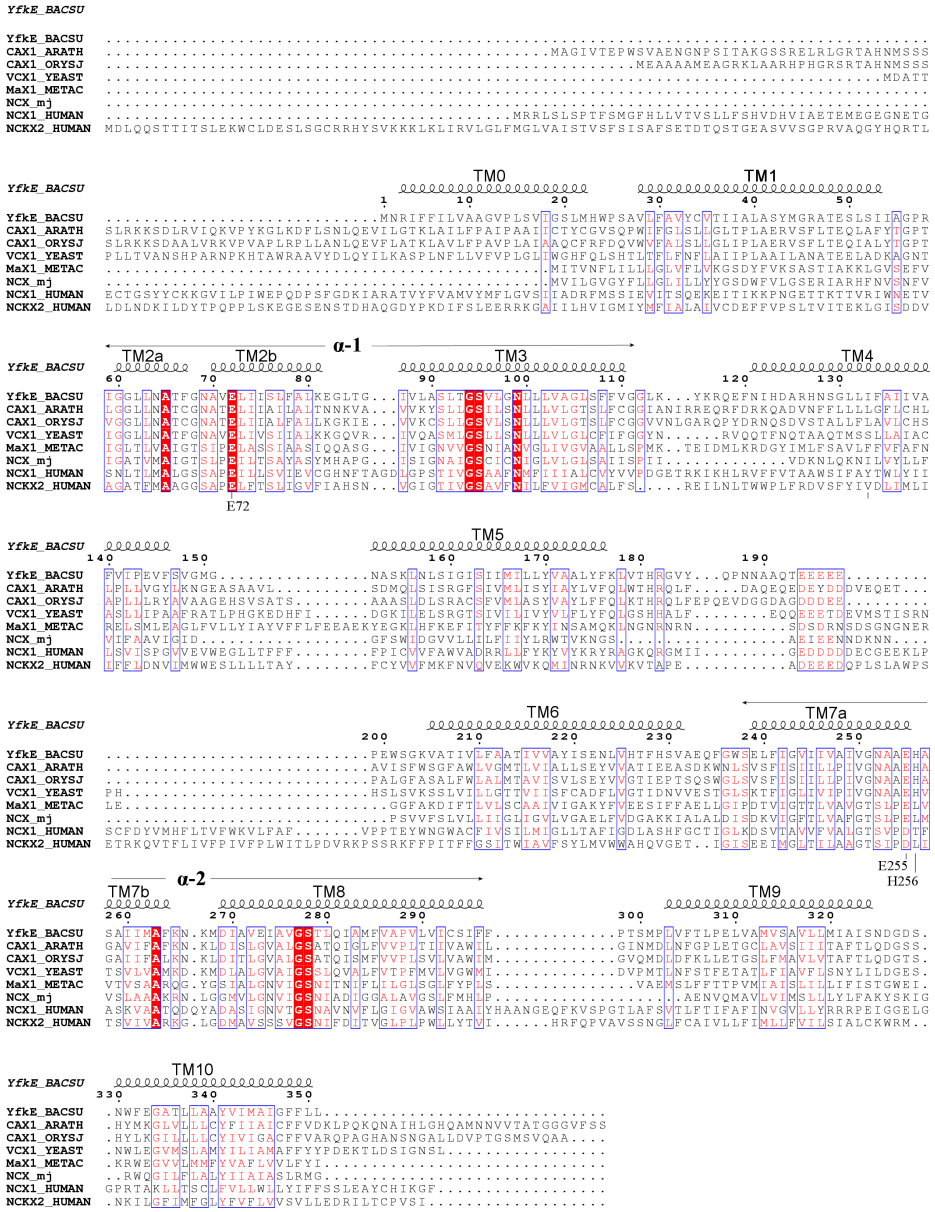

Fig. S1 | Structural-based sequence alignment of CaCA family proteins. The protein sequences are respectively from H^+/Ca^{2+} exchanger proteins: *Bacillus subtilis* (YfkE_BACSU), *Arabidopsis thaliana* (CAX1_ARATH), *Oryza sativa subsp. japonica* (Rice, CAX1A_ORYSJ), *Saccharomyces cerevisiae* (VCX1_YEAST); Na^+/Ca^{2+} exchanger proteins: *Methanosarcina acetivorans* (MaX1_METAC), *Methanocaldococcus jannaschii* (NCX_Mj), Na^+/Ca^{2+} exchanger from human Na^+/Ca^{2+} exchanger 1 (NCX1_HUMAN, truncated from residues 277 to 756) and human K^+-dependent Na^+/Ca^{2+} exchanger 2 (NCKX2_HUMAN, truncated from residues 317 to 450). As indicated in parentheses, the large intracellular loops between TMs 5-6 are not subjected to alignment. The sequence numbering is for YfkE. Two conserved α-repeat regions (α-1 and α-2) are highlighted. The sequence alignment was generated using the program ClustalW (12) and the figure was prepared by the program ESPript (13).

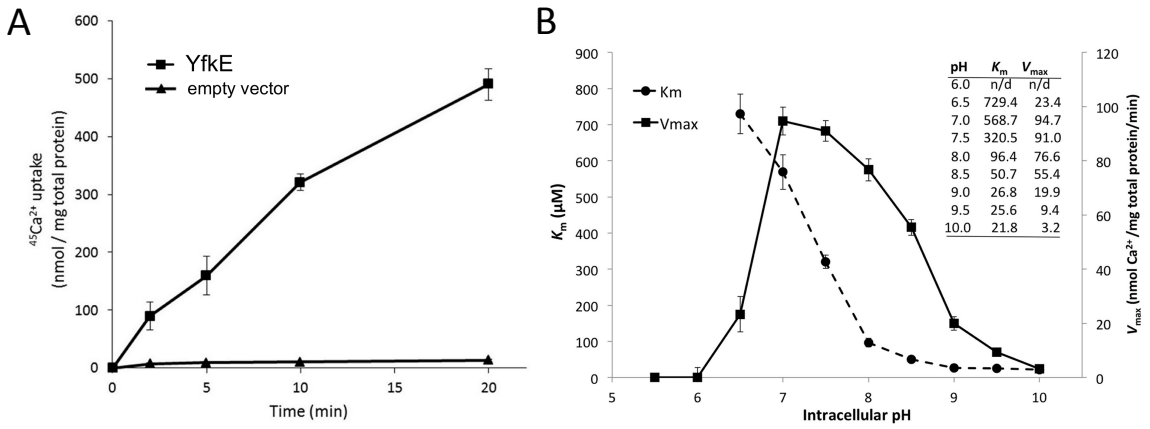

Fig. S2 | Ca^{2+} transport activity of YfkE and its pH dependence. (*A*) Time-dependent Ca^{2+} transport assay in membrane vesicles made from cells expressing YfkE (*squares*) or the empty vector (*triangles*). (*B*) Michaelis-Menten kinetic analysis of the YfkE wild type protein showing pH dependence of the K_m (circle) and V_{max} (square) values.

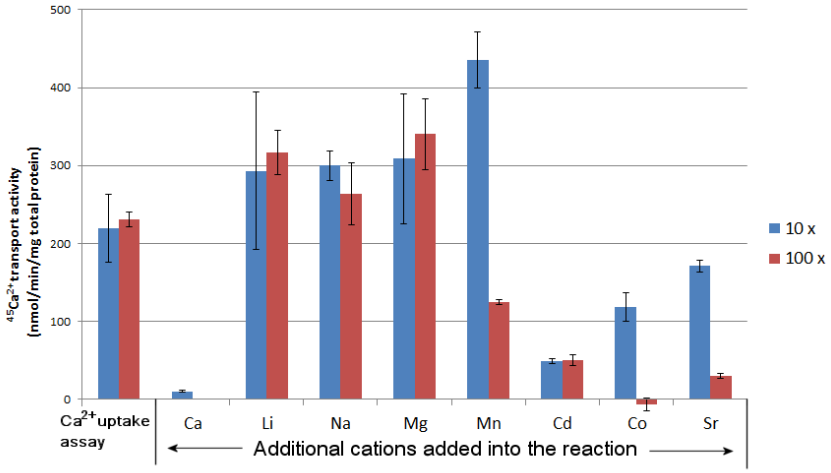

Fig. S3 | Cation inhibitory assay of YfkE. $^{45}Ca^{2+}$ uptake assays of YfkE were measured in the presence of 10 fold (*blue* bars) or 100 fold (*red* bars) higher concentration of other cations. The results were normalized with the empty vector control. The $^{45}Ca^{2+}$ transport activity of YfkE is inhibited much stronger by additional cold Ca^{2+} than Cd^{2+}, Co^{2+}, Sr^{2+} or Mn^{2+} (only at higher concentration). No inhibition was observed by addition of Li^{+}, Na^{+} or Mg^{2+}. The modest activations by Li^{+}, Na^{+} or Mg^{2+} are perhaps caused by non-specific ion electrostatic interaction with the protein due to their similar activation and their independence of cation concentration.

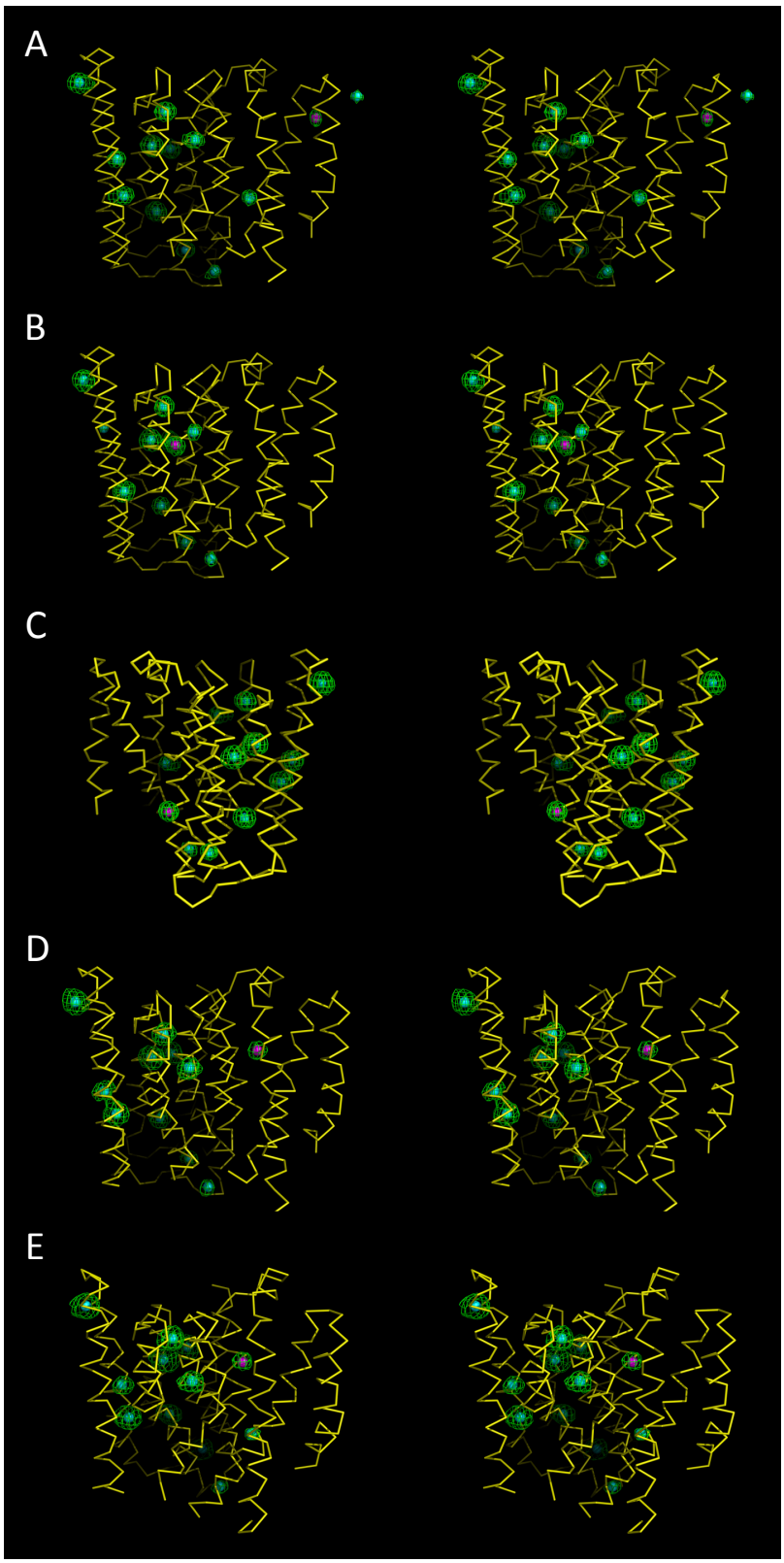
A
B
C
D
E

Fig. S4 | Experimental electron density maps of the Se heavy atoms in the YfkE mutant structures. Stereo views of the backbone of one YfkE molecule in *yellow* ribbons superimposed with an anomalous Fourier map calculated from the SeMet dataset of five methionine mutants: (*A*) L29M; (*B*) L92M; (*C*) L106M; (*D*) L224M; (*E*) V247M. The maps are displayed as *green* meshes contoured at 5 σ (*A-D*) or 4 σ (*E*). Identified Se heavy atom positions are showed as *Cyan* spheres. New Se sites introduced by methionine mutations are showed as *Magenta* spheres.

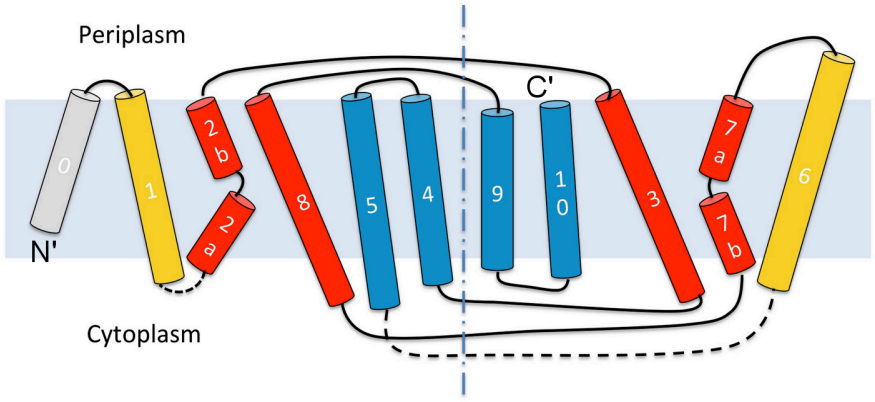

Fig. S5 | Topology of a YfkE protomer. Topology and structural organization of TM helices showing anti-parallel pseudo twofold symmetry formed by trimer-forming TMs 4-5 & 9-10 (*blue*), α-repeat TMs 2-3 & 7-8 (*red*), exterior TMs 1 & 6 (*yellow*) and distal TM0 (*grey*). The central *blue* dashed line represents the pseudo twofold symmetry. Two cytoplasmic loops unresolved in the structure are depicted as *black* dashed line: 1) between TMs 1 and 2a, 2) between TMs 5 and 6.

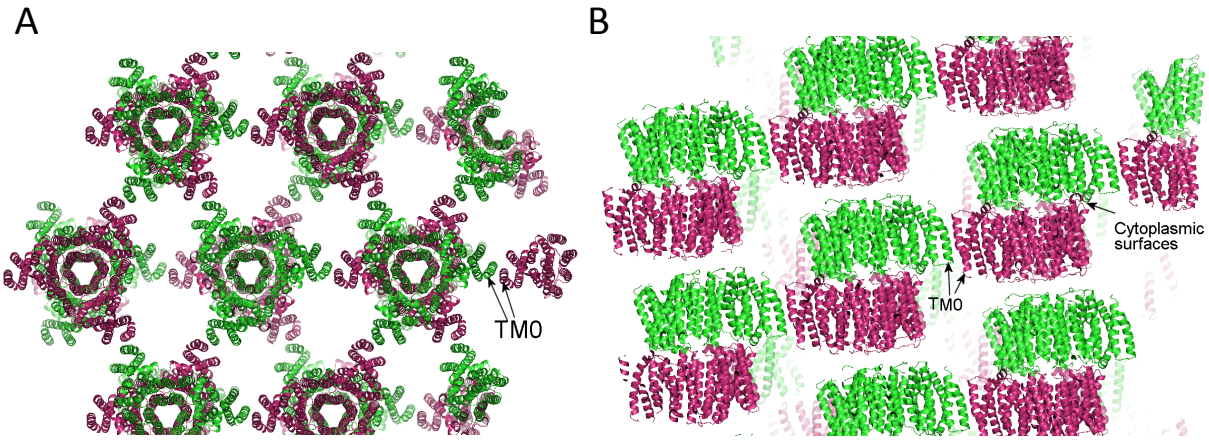

Fig. S6 | Molecular packing in the YfkE protein crystal lattice. (*A*) A top view. *(B)* A side view from the membrane plane. Two anti-parallel YfkE trimers colored in *red* or *green* stack together with their large cytoplasmic surfaces facing each other. TM0 helices from two adjacent YfkE trimers form weak anti-parallel interaction.

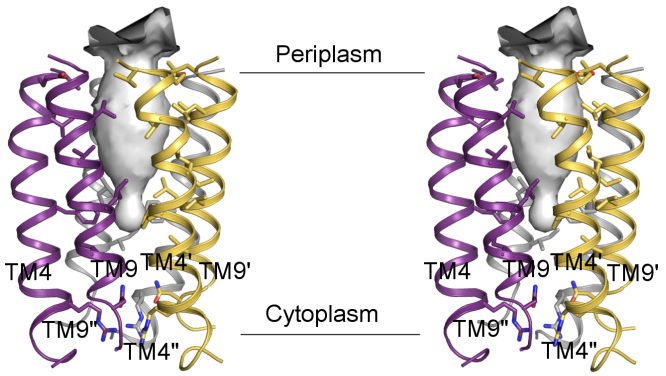

Fig. S7 | Stereo view of the central cavity of a YfkE trimer. Three copies of TMs 4 and 9 (in *grey, yellow* or *purple)* from a YfkE trimer form a vase-like hydrophobic cavity opening on the extracellular surface along the threefold axis. Residues forming the cavity are depicted as sticks with their N/O atoms in *blue/red*. The surface representation of the cavity was calculated by *MOLE* (14).

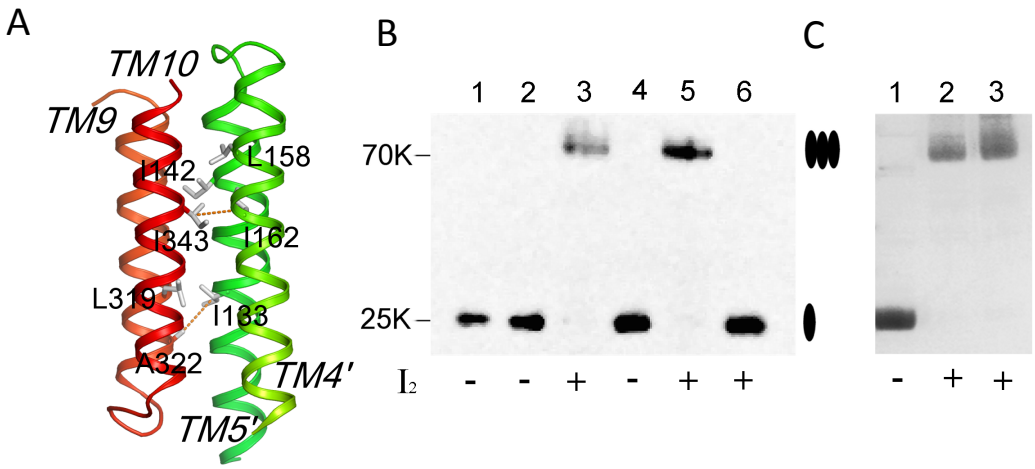

Fig. S8 | Disulfide crosslinking on the trimeric interface. (*A*) YfkE trimer interface formed by hydrophobic residues from TMs 4'-5' of one protomer (*green*) and from TMs 9-10 of its adjacent protomer (*red*). Two pairs of residues (I133/A322 and I162/I343) were chosen for disulfide crosslinking based on their proximal localizations on the trimer interface. The distance of their C alpha atoms in the pair is 6.1 Å (I133/A322) or 7.3 Å (I162/I343), which are similar to the distance of two cysteine residues forming a disulfide bond (6.5-7.2 Å) in antibody structure (PDB: 1UWG) (15). The disulfide crosslinking is mimicked with the dashed lines. (*B*) Western blot analysis of unpurified crosslinked YfkE mutant proteins on the crude *E. coli* membranes before (-) and after (+) incubation with 0.5 mM iodine. **1**: YfkE wild type; **2, 3**: I133C/A322C; **4, 5**: I162C/I343C; **6**: I133C/I343C. (*C*) Disulfide crosslinking of detergent-purified proteins visualized by commassie staining. **1**: YfkE wild type; **2**: I133C/A322C; **3**: I162C/I343C.

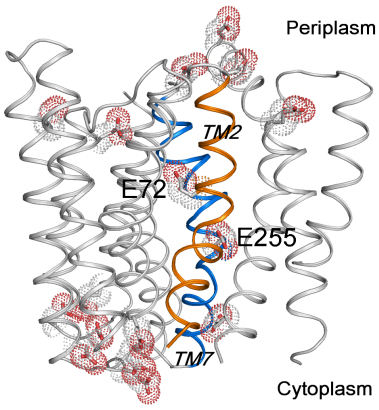

Fig. S9 | Localization of carboxylate residues in a YfkE protomer. Each YfkE protomer (cartoon) has 16 Asp/Glu residues (sticks and dots). The residues E72 and E255 from TMs 2 (*orange*) and 7 (*blue*) are embedded in the membrane in contrast to other carboxylate residues located on the periplasmic or cytoplasmic surface.

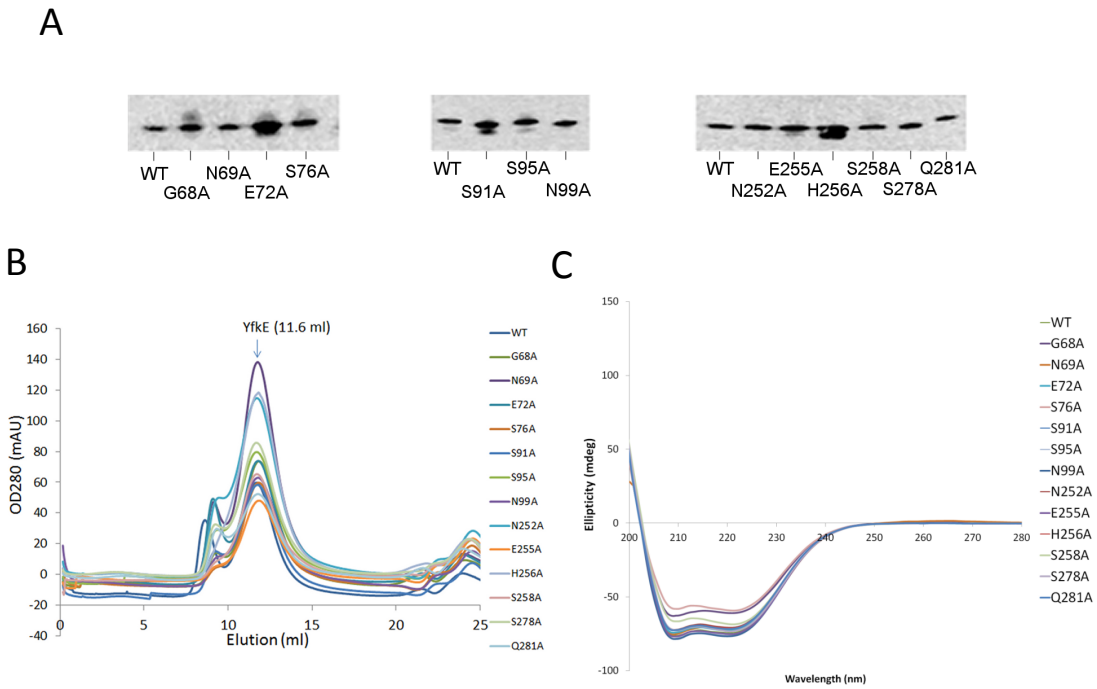

Fig. S10 | Protein characterization of YfkE mutants. (*A*) Western blot analysis of YfkE wild type or mutants showing their comparable expression yields on the everted vesicles. Western blot was performed with same amount of the vesicles in each sample using an anti-His tag antibody HRP conjugate. (*B*) Size-exclusion chromatography of the purified wild type and mutant proteins eluted from Superdex 200 10/300 GL column indicating the similarity of their protein monodispersities. The samples from Ni-NTA affinity chromatography were loaded with the same volume without adjusting protein concentration. Two protein standards, aldolase (158 KDa) and ferritin (440 KDa),

were eluted at the volume of 12.5 and 10.5 ml on this gel filtration column, respectively. Therefore, the YfkE protein may exist as an oligomer in the solution given its molecular weight of 37 kDa and estimated DDM micelle size of 70 KDa (16). (*C*) CD spectra of the purified YfkE wild type and alanine mutant proteins indicating their similar α helical conformations.

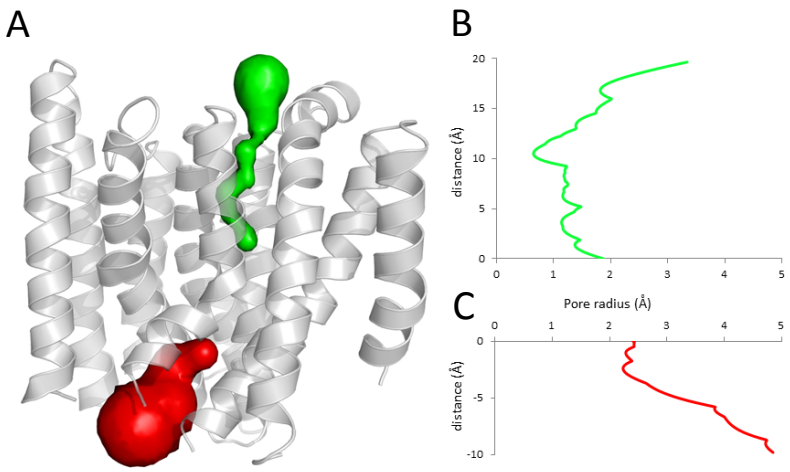

Fig. S11 | Proposed Ca^{2+} translocation pathway. (*A*) Surface representation of the pores on the periplasmic (*green*) or the cytoplasmic side (*red*) of YfkE obtained with the program *MOLE* (14). One YfkE protomer is drawn as grey cartoons viewed from the membrane plane. Their corresponding radius along the pores is displayed as (*B*) *green* curve or (*C*) *red* curve.

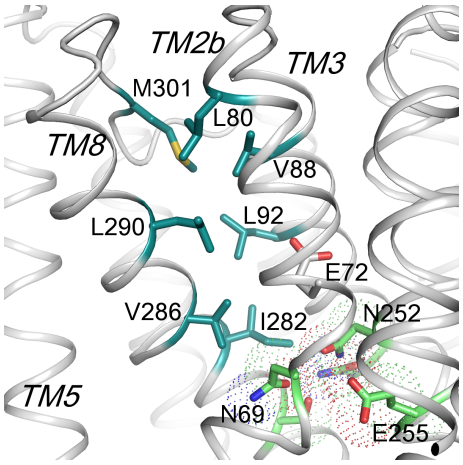

Fig. S12 | Hydrophobic interface between TMs 2b-3 and TM8. TMs 2b-3 tilt on TM8. Their TM interface formed by hydrophobic residues (*blue*). The residues at the Ca^{2+} binding site of the "umbrella" is colored in *green* and highlighted with dots. TM8 is stabilized by TM5 and the membrane-embedded loop between TMs 8 and 9. The side chain of M301 inserts deeply into the helical interface.

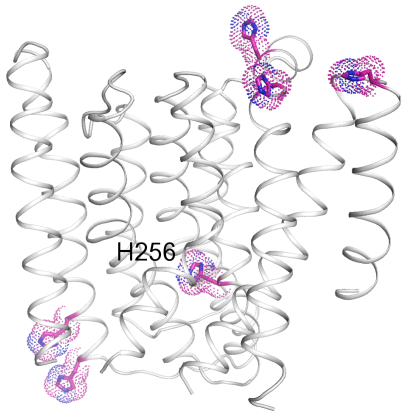

Fig. S13 | Localization of the residue H256 in a YfkE protomer. Each YfkE protomer (*grey* ribbon) has six histidine residues depicted as pink sticks and dots, among which five are located on the periplasmic or cytoplasmic surface of the molecule. H356 is the only histidine residue embedded in the membrane and along the Ca^{2+} translocation pathway.

SI Appendix, Tables

Table S1 | Statistics of data collections and structural refinements

	Se-derivative YfkE	Native YfkE
Data collection		
Wavelength (Å)	0.9790 (peak)	1.6050
Space group	*R*3	*R*3
Unit cell dimensions (Å)	170.2, 170.2, 95.2 (a, b, c) 90º, 90º, 120º (α, β, γ)	168.5, 168.5, 94.2 (a, b, c) 90º, 90º, 120º (α, β, γ)
Resolution (Å)	85-3.1 (3.27-3.1)	50-3.05 (3.21-3.05)
Number of Unique reflections	19588 (1008)	18024 (945)
I/δ	21.0 (2.1)	31.8 (4.1)
R_{merge} (%)	5.2 (74.5)	9.1 (93.2)
Redundancy	5.9 (5.8)	23.2 (22.9)
Completeness (%)	100 (100)	99.9 (100)
Refinement		
R_{work} / R_{free}	23.9% / 26.8%	21.6% / 26.1%
Number of atoms	4781	4800
Protein	4769	4782
Water	12	18
Overall *B* value (Å2)**	92	83.56
RMSD bond lengths (Å)	0.006	0.007
RMSD bond angels (°)	1.169	1.254
Ramachandran plot (%) **Favorite and allowed**	100	100

Note: Values in parentheses are for the highest resolution shell. The Se-derivative data is the L77M peak dataset in Table S2, but was processed in space group *R*3.

Table S2 | Statistics of data collections of selenomethionine-derivative YfkE crystals for phasing

	L77M			L29M	L92M	L106M	L224M	V247M
Wavelength (Å)	0.9790 (Peak)	0.9793 (Inflection)	0.9641 (Remote)	0.9788 (Peak)	0.9789 (Peak)	0.9788 (Peak)	0.9789 (Peak)	0.9794 (Peak)
Space group	$R3_2$	$R3_2$	$R3_2$	$R3_2$	$R3_2$	$R3_2$	$R3_2$	$R3_2$
Unit cell dimensions (Å) (a, b, c)	170.2, 170.2, 95.2	170.2, 170.2, 95.2	170.2, 170.2, 95.2	170.8, 170.8, 95.0	170.2, 170.2, 94.9	172.1, 172.1, 95.5	170.2, 170.2, 94.6	170.5, 170.5, 190.8
Resolution (Å)	85.1-3.1 (3.27-3.1)	84.9-3.1 (3.27-3.1)	86.3-3.1 (3.27-3.1)	85-3.3 (3.48-3.3)	85-4.0 (4.22-4.0)	86-3.7 (3.9-3.7)	95-4.0 (4.22-4.0)	116-4.0 (4.22-4.0)
Unique reflections	9734 (1413)	9663 (1350)	10159 (1467)	8136 (1190)	4562 (658)	5911 (842)	4546 (652)	9185 (1317)
Redundancy	11.2 (11.4)	11.3 (11.4)	11.3 (11.4)	11.2 (11.4)	11.1 (11.2)	11.2 (11.4)	11.2 (11.3)	39.4 (40.9)
I / δ	21.3 (2.9)	24.3 (4.2)	21.8 (2.5)	26.2 (3.0)	17.7 (2.9)	31.0 (3.5)	18.8 (2.9)	25.8 (5.7)
R_{merge} (%)	7.2 (85.9)	6.4 (59.1)	7.3 (96.6)	6.3 (76.0)	9.7 (82.2)	4.6 (68)	9.3 (82.7)	9.6 (80.8)
Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)	99.9 (100)	99.9 (99.9)	99.9 (100)	99.8 (99.8)
Anomalous Redundancy	5.7 (5.7)	5.7 (5.7)	5.8 (5.8)	5.8 (5.8)	5.8 (5.8)	5.8 (5.8)	5.8 (5.8)	20.6 (21.2)
R_{ano} (%)	4.8 (27.5)	3.9 (20.6)	3.9 (31.6)	4.8 (26.9)	6.3 (29.2)	4.8 (23.0)	6.1 (27.2)	4.7 (13.6)
Anomalous Completeness (%)	100 (100)	100 (100)	100 (100)	100 (100)	99.9 (100)	99.9 (99.9)	99.9 (100)	99.8 (99.9)

Note: Values in parentheses are for the highest resolution shell.

SI Appendix, References

1. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41: 207-234.
2. Sreenath HK *et al.* (2005) Protocols for production of selenomethionine-labeled proteins in 2-L polyethylene terephthalate bottles using auto-induction medium. *Protein Expr Purif* 40:256-267.
3. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125-132.
4. Adams PD *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-221.
5. Vonrhein C, Blanc E, Roversi P & Bricogne G (2007) Automated structure solution with autoSHARP. *Methods Mol Biol* 364:215-230.
6. Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
7. Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53: 240-255.
8. McCoy AJ *et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658-674.
9. DeLano WL (2002) The PyMOL Molecular Graphics System. www.pymol.org
10. Wu J & Kaback HR (1996) A general method for determining helix packing in membrane proteins in situ: Helices I and II are close to helix VII in the lactose permease of *Escherichia coli*. *Proc Natl Acad. Sci. USA*. **93**, 14498-14502.
11. Rosen BP & Tsuchiya T (1979) Preparation of everted membrane vesicles from *Escherichia coli* for the measurement of calcium transport. *Methods Enzymol* 56:233-241.
12. Larkin MA *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
13. Gouet P, Courcelle E, Stuart DI. & Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15: 305-308.
14. Petrek M, Kosinova P, Koca J, & Otyepka M (2007) MOLE: a Voronoi diagram-based explorer of molecular channels, pores, and tunnels. *Structure* 15(11):1357-1363.

15. Zheng L, Baumann U & Reymond JL. (2004) Molecular mechanism of enantioselective proton transfer to carbon in catalytic antibody 14D9. *Proc Natl Acad Sci USA* 101(10): 3387-3392.
16. VanAken T, *et al*. (1986) Alkyl glycoside detergents: synthesis and applications to the study of membrane proteins. *Methods Enzymol* 125: 27–35.