Supporting Information

Yu et al. 10.1073/pnas.1415037112

SI Methods

Expression and Purification of Art(QN)$_2$. The artQ and artN genes were cloned from *T. tengcongensis* genomic DNA into a pACYCDuet vector (Novagen) that can coexpress *artQ* and *artN* stimultaneously using separate T7 promoters. An engineered C-terminal 6× His tag was fused to artQ (encoding the transmembrane subunit), which was cloned into multiple cloning sites-2, whereas the artN gene (ATPase subunit) was inserted in multiple cloning sites-1. Overexpression of Art(QN)$_2$ was induced in *E. coli* strain C43 by 0.5 mM isopropyl-β-D-thiogalactoside when the cell density reached OD_{600} = 1.2. After growth for 4 h at 37 °C, the cells were collected, resuspended in buffer containing 25 mM Tris·HCl pH 8.0 and 200 mM NaCl, and lysed by sonication. Cell debris was removed by centrifugation at 15,422 × *g* for 15 min. The supernatant was subsequently subjected to ultracentrifugation at 173,021 × *g* for 1 h. The membrane fraction (pellet) was resuspended in buffer A containing 25 mM Tris·HCl pH 8.0, 20 mM imidazole, and 500 mM NaCl. After the addition of n-dodecyl-β-D maltopyranoside (DDM; Anatrace) at a final concentration of 1.5% (wt/vol), the mixture was incubated for 5 h with slow stirring at 4 °C. After another ultracentrifugation step at 173,021 × *g* for 30 min, the supernatant was collected and loaded onto Ni^{2+}-nitrilotriacetate affinity resin (Ni-NTA; Qiagen). The resin was then washed with Buffer A supplemented with 0.02% DDM. Protein was eluted from the affinity resin with Buffer A supplemented with 0.02% DDM and 200 mM imidazole-HCl pH 8.0. Protein-containing fractions were pooled, concentrated, and applied to a gel-filtration resin (Superdex-200 HR 10/30; GE Healthcare), previously equilibrated with 25 mM Tris·HCl pH 8.0, 200 mM NaCl, 5 mM DTT, and 0.2% DM. The peak fractions were collected and concentrated to 8 mg/mL for crystallization trials.

Expression and Purification of ArtI. The artI gene was cloned from *T. tengcongensis* genomic DNA into a pet21b vector (Novagen) with a C-terminal 6× His tag. Overexpression of ArtI was induced in *E. coli* strain BL21(DE3) by 0.5 mM isopropyl-β-D-thiogalactoside when the cell density reached OD_{600} = 1.2. After growth for 10 h at 23 °C, the cells were collected; resuspended in Buffer A containing 25 mM Tris·HCl pH 8.0, 500 mM NaCl, and 20 mM imidazole; and lysed by sonication. After ultracentrifugation at 13,000 × *g* for 1 h, the supernatant was loaded onto Ni-NTA resin. After a washing step using Buffer A, protein was eluted with Buffer A supplemented with 200 mM imidazole-HCl pH 8.0. Protein-containing fractions were pooled, concentrated, and subjected to gel filtration (Superdex-200; GE Healthcare) in 25 mM Tris pH 8.0, 200 mM NaCl, and 5 mM DTT. The peak fractions were collected and concentrated to 15 mg/mL for crystallization trials.

Denaturation and Renaturation of ArtI. For ATPase assays, ArtI was subjected to a denaturation/renaturation procedure to remove prebound substrate (1). In brief, 10 mL of ArtI (1 mg/mL) was denatured by adding guanidinium hydrochloride to a final concentration of 6 M and subsequently dialyzed against 1L of buffer containing 25 mM Tris·HCl pH 8.0, 200 mM NaCl, and 6 M guanidinium hydrochloride for three times at an interval of 12 hours. Renaturation was achieved by dialysis against 1L of buffer containing 25 mM Tris·HCl pH 8.0 and 200 mM NaCl for five times at an interval of 12 hours. Subsequently, the dialysate was ultracentrifuged at 220,000 × *g* for 15 min to remove aggregated protein and concentrated by a centrifugal device (Amicon; EMD Millipore). Glycerol and DTT were added to concentrations of 10% (vol/vol) and 10 mM, respectively.

Preparation of Proteoliposomes. Incorporation of Art(QN)$_2$ variants into liposomes prepared from *G. stearothermophilus* total lipids was carried out as described previously (2). In brief, lipids (20 mg) were dried under a stream of nitrogen, redissolved in 1 mL of 50 mM Mops-KOH pH 7.5, containing 1% n-Octyl-β-D-glucopyranoside (β-OG; Anatrace), and sonicated for 15 min. Subsequently, Art(QN)$_2$ variants (50 μg) were added to 125 μL of the lipid–detergent mixture, resulting in a final volume of 300 μL. Proteoliposomes were formed by removal of detergent by adsorption to Biobeads (100 mg; BioRad) at 4 °C overnight. After the beads were replaced with a new batch, incubation continued for another 2 h. The mixture was centrifuged for 1 min at 10,000 × *g* to pellet the beads. Then proteoliposomes were recovered by ultracentifugation for 30 min at 220,000 × *g*, resuspended in 50 mM Mops-KOH pH 7.5, and assayed for ATPase activity.

ATPase Assay. ATPase activity was assayed essentially as described previously (2). Reactions were started by adding 3 mM $MgCl_2$ and 2 mM ATP to preheated (70 °C) proteoliposomes in the presence of 50 mM Mops (pH 7.5), ArtI (35 μM), L-arginine or L-histidine (100 μM each), and *ortho*-vanadate (1 mM) (where indicated). Aliquots (25 μL containing 3 μg of protein) were taken in 2-min intervals and placed into wells of a microtiter plate containing 25 μL of a 12% (wt/vol) SDS solution. The amount of liberated phosphate was determined colorimetrically with ammonium molybdate complexes using Na_2HPO_4 as the standard.

ITC Measurements. Calorimetric experiments were conducted at 25 °C with a MicroCal iTC200 instrument (GE Healthcare). The protein sample (ArtI, 22–260 aa) was dialyzed against the buffer containing 20 mM Hepes pH 7.5 and 200 mM NaCl. Arginine and other amino acids were prepared in the same buffer and the pH was checked before use. The data were fitted using the Origin 7 software package (Microcal).

Crystallization, Data Collection, and Structure Determination. Crystals were grown at 18 °C by the hanging-drop vapor diffusion method. Crystals of Art(QN)$_2$ complex appeared overnight in the well buffer containing 12% (wt/vol) PEG 4000, 0.1 M sodium chloride, 0.1 M sodium citrate tribasic dehydrate pH 5.0, 0.1 M magnesium chloride, and 8 mM 3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (CHAPSO), typically growing to full size in ∼3–4 d. Before data collection, crystals were transferred into a solution containing 2.5 mM ATP, 18% (vol/vol) glycerol, 12% (wt/vol) PEG 4000, 0.1 M sodium chloride, 0.1 M codium citrate tribasic dehydrate pH 5.0, 0.1 M magnesium chloride and 8 mM CHAPSO. Crystals of ArtI were grown in the well buffer containing 0.2 M sodium chloride, 0.1 M Bis-Tris pH 6.5, and 25% (wt/vol) PEG 3350.

All of the data were collected at the Shanghai Synchrotron Radiation Facility Beamline BL17U and integrated and scaled using the HKL2000 package (3). Further processing was carried out using programs from the CCP4 suite (4). Data collection statistics are summarized in Table S1. The molecular replacement was performed with PHASER (5) using *E. coli* Methionine ABC Transporter (PDB ID code 3DHW) as a model. The final model rebuilding was performed using Coot (6), and the protein structure was refined with phenix (7) using noncrystallographic symmetry and stereochemistry information as restraints. The structure

of ArtI was solved with molecular replacement using amino acid ABC transporter substrate-binding protein AbpA from *Streptococcus pneumoniae* (PDB ID code 4I62) as a model, and structural figures were generated in PyMOL (8).

1. Hall JA, Davidson AL, Nikaido H (1998) Preparation and reconstitution of membrane-associated maltose transporter complex of *Escherichia coli*. *Methods Enzymol* 292:20–29.
2. Weidlich D, et al. (2013) Residues of a proposed gate region in type I ATP-binding cassette import systems are crucial for function as revealed by mutational analysis. *Biochim Biophys Acta* 1828(9):2164–2172.
3. Zbyszek Otwinowski WM, Charles Carter W, Jr (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
4. Collaborative Computational Project No. 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–763.
5. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Cryst* 40(Pt 4):658–674.
6. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
7. Adams PD, et al. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58(Pt 11):1948–1954.
8. DeLano WL (2002) *The PyMOL Molecular Graphics System* (Delano Scientific, San Carlos, CA).
9. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21): 2947–2948.

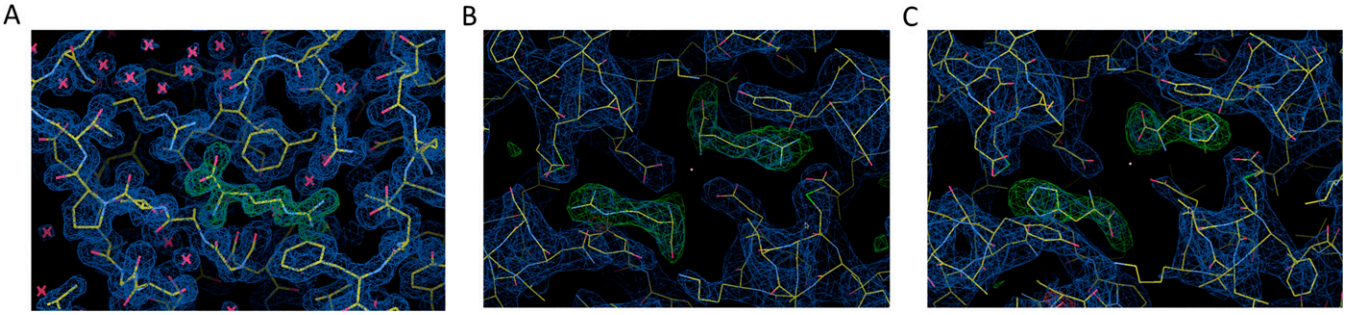


Fig. S1. Representative views of the electron density of ArtI in complex with Arg and Art(QN)$_2$ in complex with substrates. (*A*) The 2Fo-Fc electron density map, shown in blue mesh, is contoured at 1.5σ, and the Fo-Fc electron density map, shown in green mesh, is contoured at 5σ. (*B*) Electron density for the bound Arg. The 2Fo-Fc electron density map, shown in blue mesh, is contoured at 1.5σ, and the Fo-Fc electron density map, shown in green mesh, is contoured at 3.3σ. (*C*) Electron density for the bound His. The 2Fo-Fc electron density map, shown in blue mesh, is contoured at 1.5σ, and the Fo-Fc electron density map, shown in green mesh, is contoured at 3.5σ.

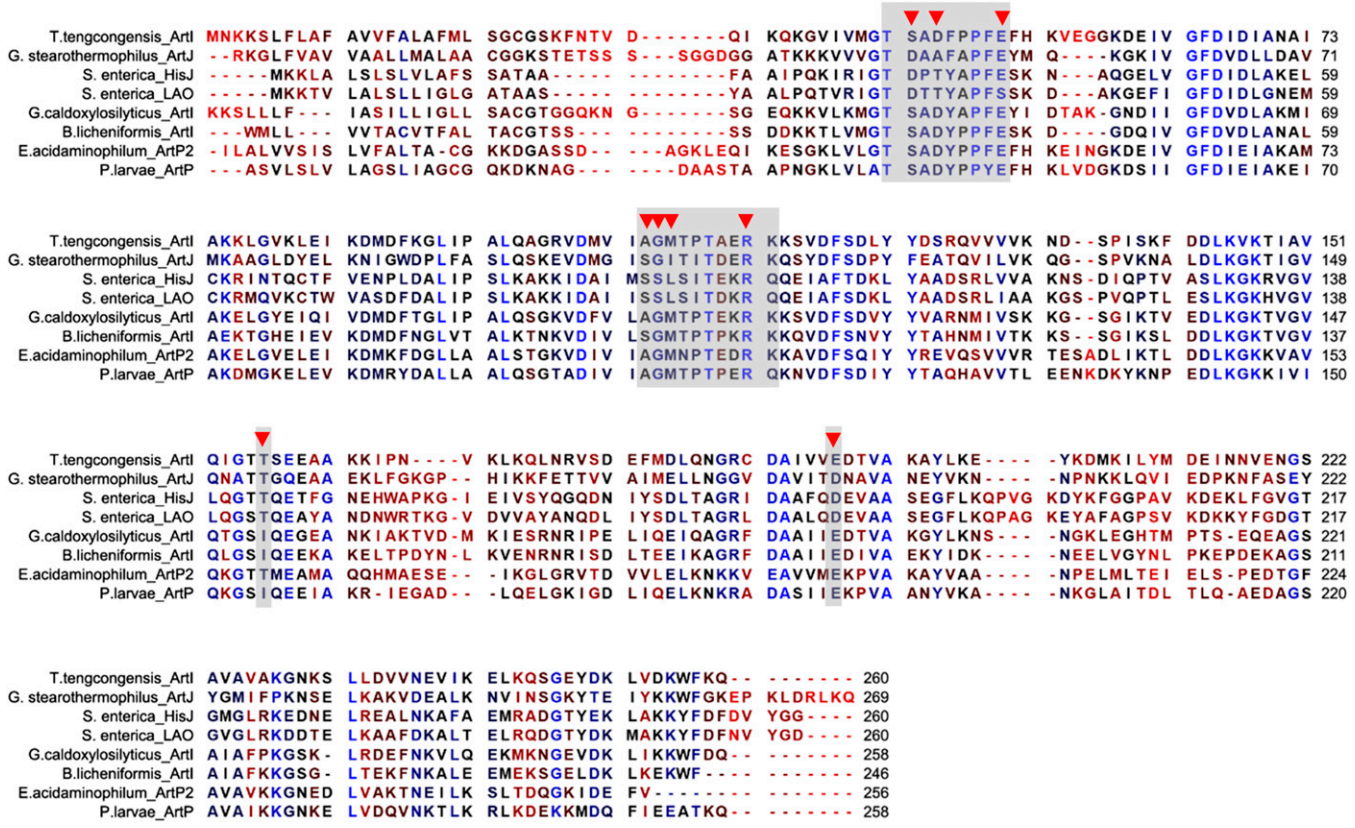


Fig. S2. The residues participating in substrate binding are conserved in ArtI homologs from different species. Shown is the sequence alignment of ArtI homologs from representative organisms. Conserved residues are highlighted in blue. The key residues involving in substrate recognition are indicated by red triangles. Sequence alignment was carried out using ClustalW (9).

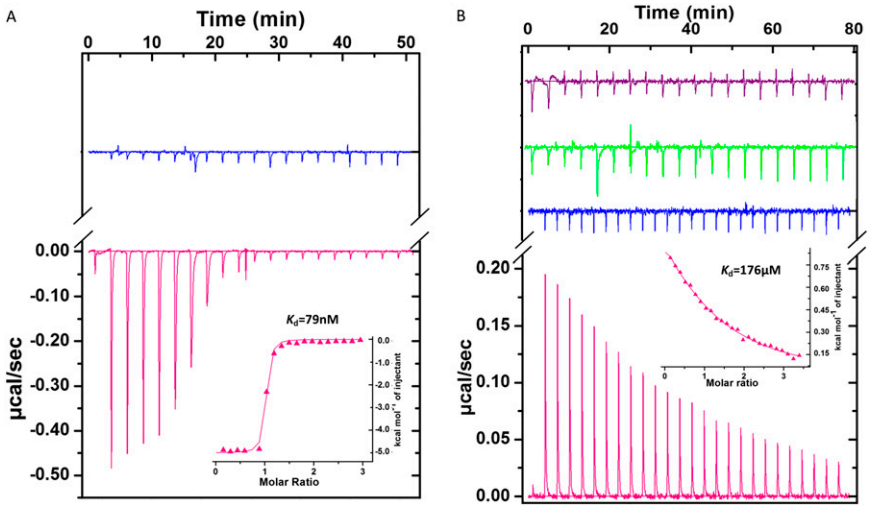


Fig. S3. ArtI displays different binding affinities for Arg and His. (*A*) ITC curves of Arg titrated into unloaded ArtI. Pink trace, 500 µM arginine titrated into 35 µM protein; blue trace, 500 µM arginine titrated into buffer. The K_d value of arginine titrated into buffer was not detectable. The first peak in the thermogram was not used for analysis. (*B*) ITC curves of histidine, glutamate, and threonine titrated into unloaded ArtI. Pink trace, 2 mM histidine titrated into 120 µM protein; blue trace, 2 mM histidine titrated into buffer; green trace, 1 mM glutamate titrated into 100 µM protein; purple trace, 1 mM of threonine titrated into 100 µM protein. K_d values of glutamate and threonine titrated into unloaded ArtI, as well as histidine titrated into buffer, were not detectable. The first peak in the thermogram was not used for analysis.

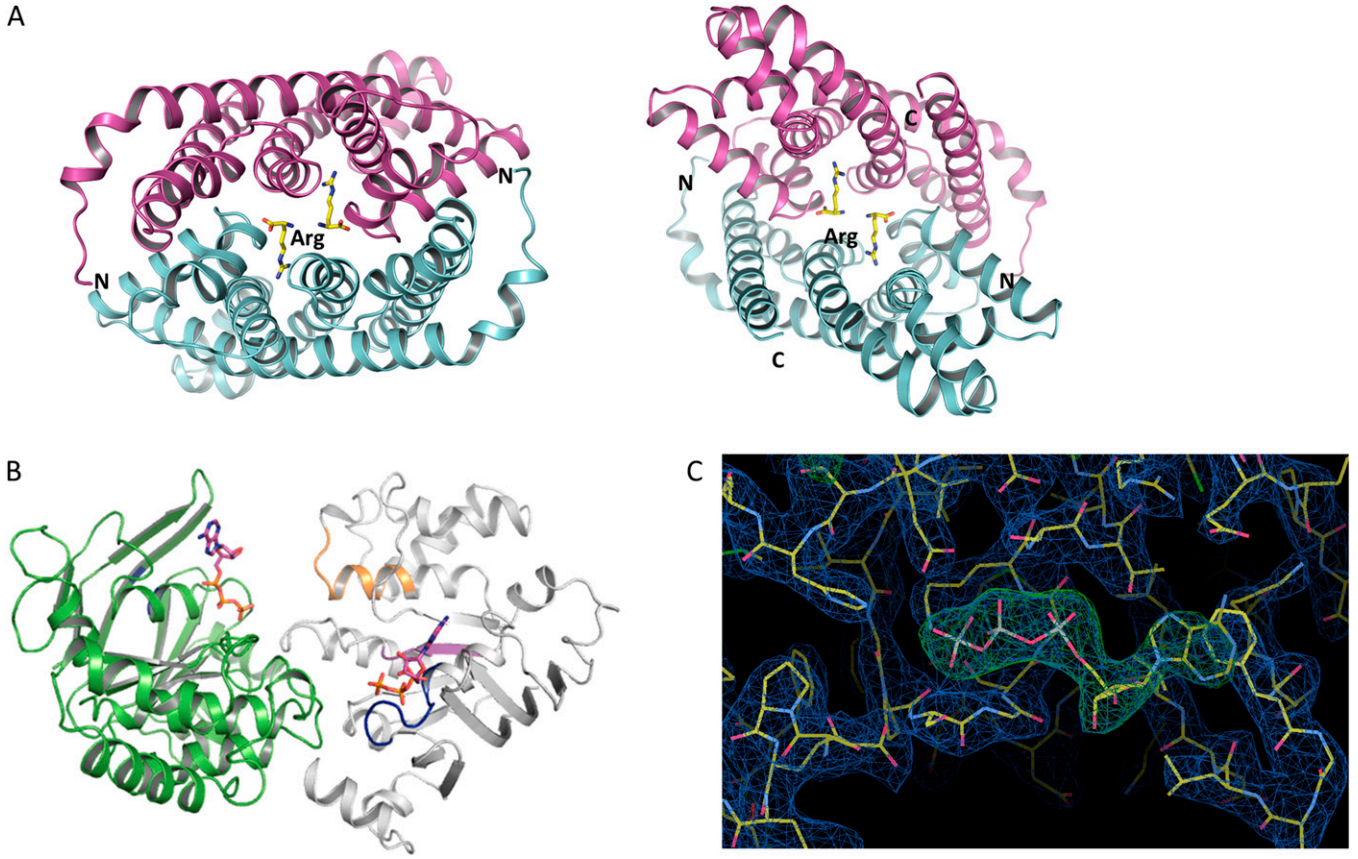


Fig. S4. Art(QN)$_2$ binds to two substrates and two ATPs. (*A*) Ribbon diagram of TM helices surrounding two substrate binding sites viewed along the membrane bilayer from the extracellular side (left) or from cytoplasmic side (right). ArtQ subunits are colored in cyan and magenta. Two Arg molecules are shown in ball-and-stick models. (*B*) Overall structure of the ArtN dimer in complex with ATPs. The Walker A motif (residues 34–41) is colored blue, the Walker B motif (residues 157–162) is colored magenta, and the ABC signature motif (residues 137–146) is colored orange, ArtN subunits are colored gray and green, respectively. Two ATP molecules are shown in ball-and-stick models. (*C*) Representative view of the electron density of ATP-bound Art(QN)$_2$. The 2Fo-Fc electron density map, shown in blue mesh, is contoured at 1.5σ, and the Fo-Fc electron density map, shown in green mesh, is contoured at 3σ.

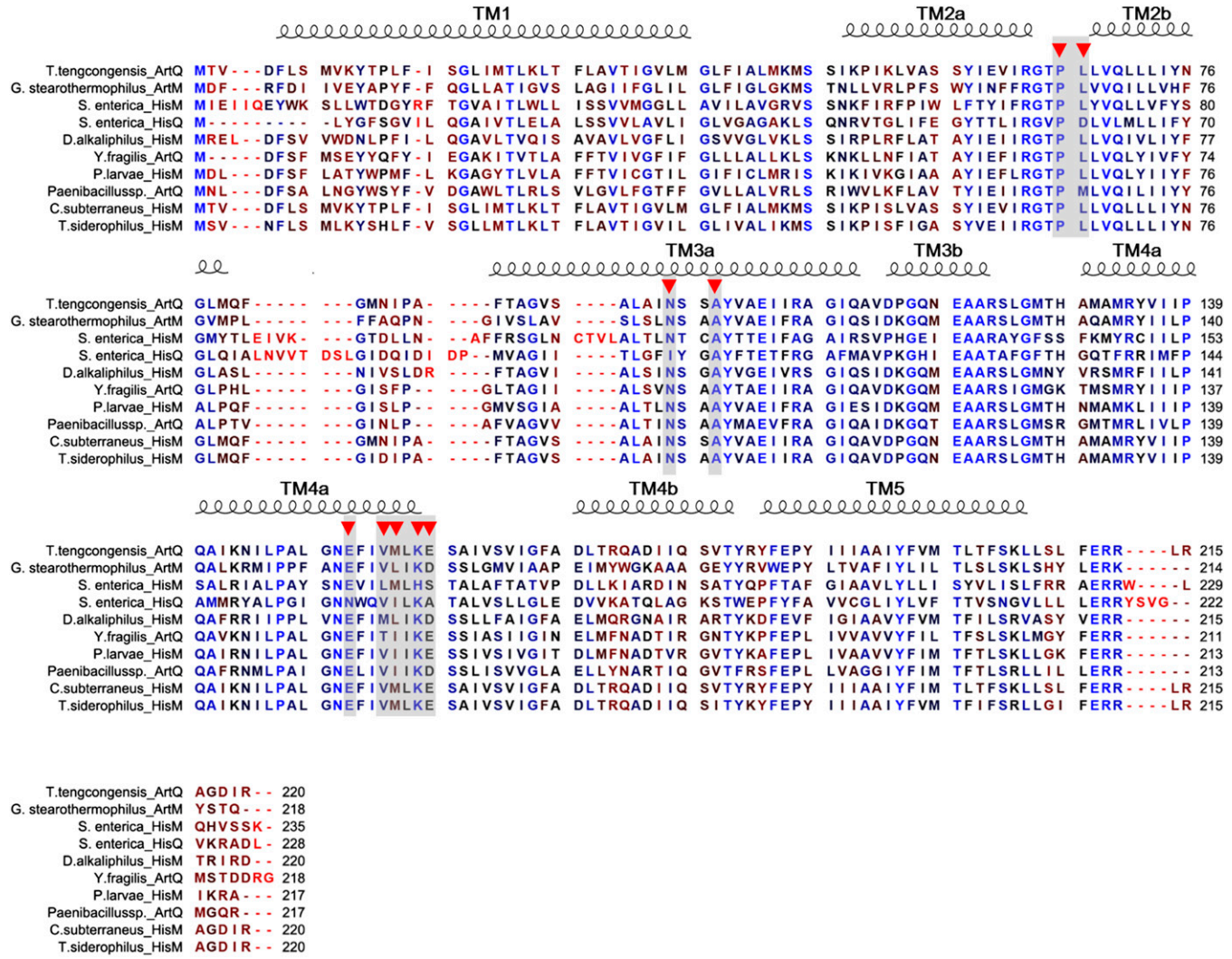


Fig. S5. The residues involved in substrate binding are conserved in ArtQ homologs from different species. Sequence alignment of ArtQ homologs from representative organisms. ArtQ^{P66}, ArtQ^{L67}, ArtQ^{N98}, ArtQ^{Y102}, ArtQ^{E152}, ArtQ^{V155}, ArtQ^{E159}, and ArtQ ^{M156} residues involved in interacting with substrate are highly conserved. Those residues are highlighted in blue and indicated by red triangles. Invariant amino acids are highlighted in blue. Sequence alignment was carried out using Clustal W (9). Secondary structural elements of ArtQ are indicated.

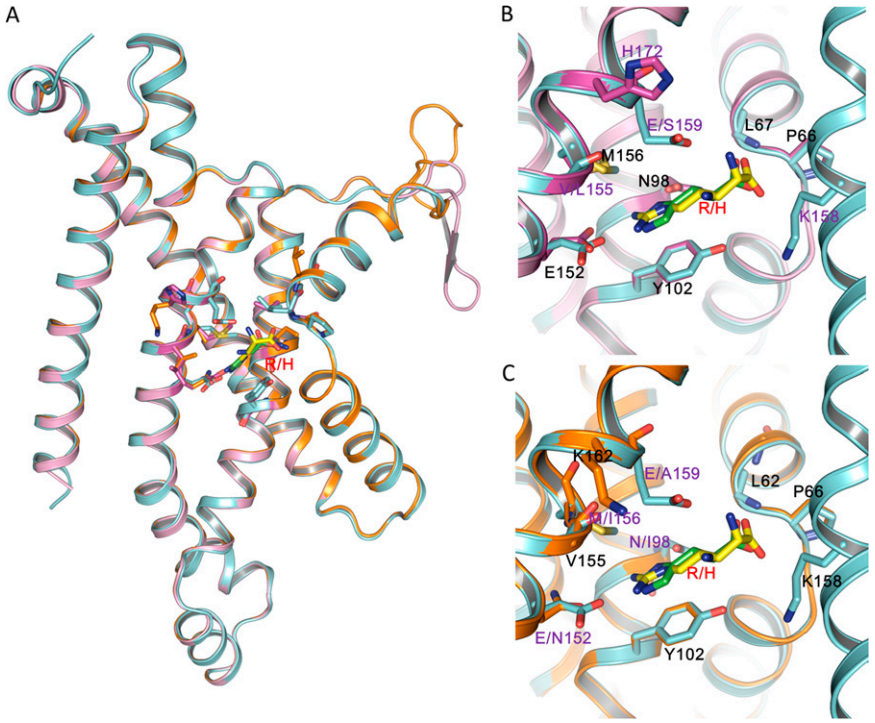
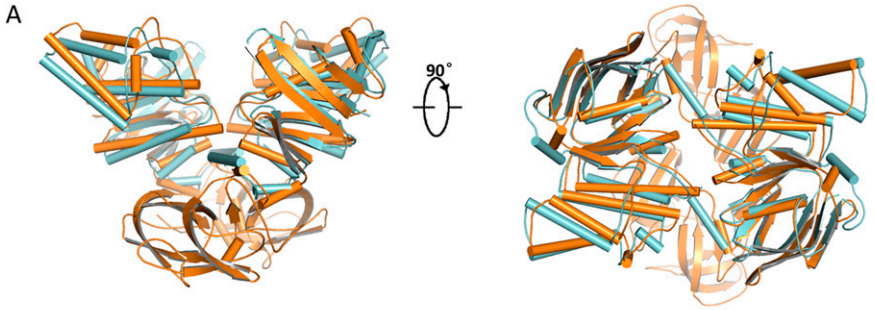
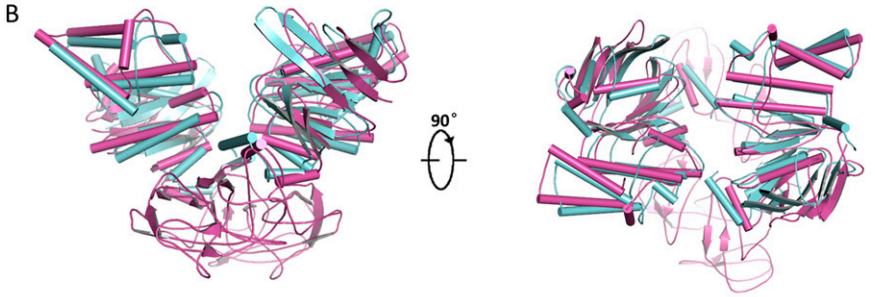


Fig. S6. The residues of ArtQ involved in substrate binding are conserved in the HisM/HisQ subunits of the His-, Lys-, Arg-transporter from *Salmonella enterica* serovar Typhimurium (*S*. Typhimurium). (*A*) Three-dimensional structure of HisM/HisQ from *S*. Typhimurium based on the structure of ArtQ. Superposition of ArtQ (cyan), HisM (pink), and HisQ (orange). Representative models of the residues from the substrate-binding site as well as Arg (R) /His (H) are shown in ball-and-stick models. (*B*) Stereoview of the structural alignment between ArtQ (cyan) and HisM (pink). In ArtQ and HisM, residues that mediate substrate binding are shown as cyan and pink sticks, respectively. Invariant and variant residues are indicated in black and magenta, respectively. (*C*) Stereoview of the structural alignment between ArtQ (cyan) and HisQ (orange). In ArtQ and HisQ, residues that mediate substrate binding are shown as cyan and orange sticks, respectively. Invariant and variant residues are indicated in black and magenta, respectively.



Superposition of ArtN (cycan) and MalK (orange) in semi-open state.



Superposition of ArtN (cycan) and MalK (magenta) in open state.

Fig. S7. The ArtN dimer with two bound ATPs is in "semi-open" state. (*A*) Stereo views of the structural alignment between the ArtN dimer and the semi-open state of the MalK dimer of the maltose transport system from *E. coli* (PDB ID code 3PV0). The MalK dimer is in orange, and the ArtQ dimmer is in cyan. (*B*) Stereo views of the structural alignment between the ArtN dimer and the open state of the MalK dimer (PDB ID code 3FH6). The MalK dimer is in magenta, and the ArtQ dimer is in cyan.

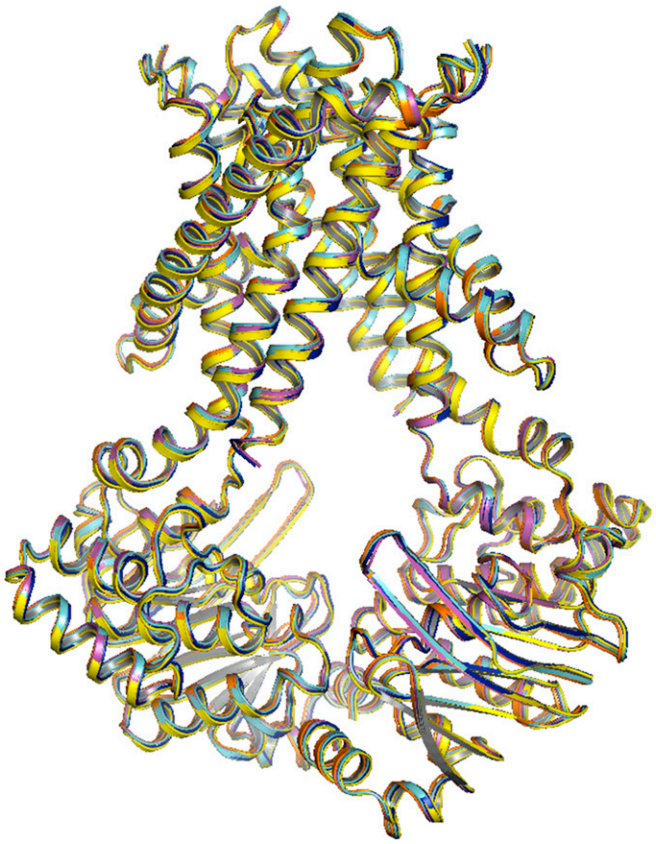


Fig. S8. No obvious changes are seen in apo-, ATP-, Arg-, His-, and Arg/ATP-bound states of Art(QN)$_2$. Shown is structural alignment among apo-bound (purple), ATP-bound (blue), Arg-bound (yellow), His-bound (orange), and Arg-/ATP-bound (cyan) states of Art(QN)$_2$.

Table S1. Data collection and refinement statistics

Type	Apo-bound Art(QN)$_2$	Arg-bound Art(QN)$_2$	Arg/ATP-bound Art(QN)$_2$	ATP-bound Art(QN)$_2$	His-bound Art(QN)$_2$	Arg-bound ArtI
Data collection statistics (beamline: SSRF-BL17U; detector: ADSC Quantum 315r; software: HKL2000)						
Space group	C2221	C2221	C2221	C2221	C2221	P1
Unit cell						
a, b, c, Å	96.096, 114.209, 302.019	96.379, 114.779, 300.202	96.765, 114.098, 298.636	96.527, 115.257, 301.546	96.247, 115.264, 301.185	39.026, 41.409, 75.604
α, β, γ, °	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	103.73, 102.72, 90.12
Wavelength, Å	0.9793	0.9793	0.9793	0.9793	0.9793	0.9789
Resolution, Å	41.54–2.8 (2.9–2.8)	49.8–2.59 (2.68–2.59)	49.50–2.50 (2.59–2.50)	43.5–3.00 (3.11–3.00)	34.72–2.80 (2.90–2.80)	30.68–1.48 (1.53–1.48)
R_{merge}, %	11.3 (65.0)	9.2 (63.6)	8.9 (82.4)	10.2 (55.4)	9.5 (49.0)	4.1 (12.0)
I/sigma	15.9 (2.14)	34.7 (3.0)	19.2 (1.71)	22.3 (2.2)	22.0 (2.1)	28.0 (10.3)
Completeness, %	95.59 (79.30)	99.8 (100)	99.3 (94.8)	98.7 (91.6)	92.9 (74.4)	95.9 (85.9)
Redundancy	6.7 (5.4)	6.8 (6.2)	5.9 (4.9)	5.5 (4.6)	4.3 (3.6)	2.2 (2.1)
Structure refinement statistics (software: phenix.refine)						
R_{work} /R_{free}	22.8/27.6	22.2/26.5	22.4/26.0	20.6/25.3	22.6/27.1	15.5/18.3
No. of atoms						
All	7,046	7,086	7,134	7,108	7,068	4,508
Side chains	3,410	3,438	3,426	3,410	3,424	1,845
Main chains	3,636	3,648	3,644	3,636	3,644	1,916
Macromolecule	7,046	7,086	7,070	7,046	7,068	3,761
Other entities	0	65	64	62	0	747
Average B-factor, Å^2						
All	92.7	73.2	88.1	93.0	86.5	12.8
Side chains	93.6	73.9	88.9	93.3	87.2	12.2
Main chains	91.9	72.5	87.4	92.0	85.9	8.0
Macromolecule	92.7	73.2	88.1	92.6	86.5	10.0
Other entities		73.2	85.5	132.2		27.0
Rmsd						
Bond lengths, Å	0.010	0.01	0.009	0.010	0.009	0.006
Bond angles, °	1.33	1.26	1.28	1.29	1.24	1.003
Ramachandran						
Favored, %	94.0	96.5	97.0	96.23	96.67	98.52
Outliers, %	0.11	0.22	0.00	0.00	0.00	0.52
PDB ID codes	4YMS	4YMT	4YMU	4YMV	4YMW	4YMX

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).