Supplementary information, Data S2 Additional results and discussion

Overall structure

The average length of the four MdfA cavity-helices is 27.0 (±1.4) residues; that of the rocker-helices is 27.3 (±1.9); and that of the support-helices is 22.3 (±2.1). The fact that the support-helices are of shorter length is consistent with the rocker-switch mechanism, in which the support-helices move less than the other two groups of TM helices (relative to the lipid bilayer) during the conformational changes. Also, the support-helices appear straighter than others, in agreement with their shorter and more uniform O_i–N_{i+4} hydrogen-bonds (H-bonds) of an average length (±STD) of 3.1 (±0.3) Å, compared to 3.3 (±0.7) Å for the cavity-helices and 3.3 (±0.6) Å for the rocker-helices.

The H-bond network in TM1 of the MdfA N-domain clearly shows irregularities. For example, the main-chain carbonyl group of G32$^{TM1.3}$ forms a H-bond with the side-chain guanidinium group of R112$^{TM4.2}$ of the motif-B, rather than with the expected main-chain amid group of I36$^{TM1.2}$. Related to this irregularity, the main-chain carbonyl group of T29$^{TM1.3}$ forms a H-bond with the side-chain hydroxyl group of Y61$^{TM2.2}$ which is also associated with the motif-B (see below). Similarly, Y323$^{TM10.3}$ within the C-domain contributes its side-chain hydroxyl group to bridge a broken main-chain H-bond between V231^{TM7,4} and L235$^{TM7.3}$ which is caused by P234$^{TM7.3}$. The cavity-helices of MFS transporters (*i.e.* TMs 1, 4, 7, and 10) usually show more irregularity in their H-bonds than the other two groups of TM helices, a fact that is probably related to their roles in substrate binding [6, 7].

Besides ligands bound to the cavity, a LDAO molecule was clearly seen in both the Dxc- and LDAO-complex crystal structures (but not in the Cm-complex structure). This LDAO molecule binds to the surface of TM region of one MdfA molecule, using its hydrophobic lipid chain. It also binds to the side chains of

E207 and D211 of another MdfA molecule, using its positively charge head group. Thus, this LDAO molecule is likely to play a role in stabilizing the crystal packing.

Structure comparison of MdfA with YajR and modelling of MdfA C_{Out} structure

MdfA and YajR (PDB ID: 3WDO) are the only two motif-B containing antiporters that have their crystal structures reported in C_{In} and C_{Out} states [6], respectively. The residues associated with motif-B appear similar in the crystal structures of YajR C_{Out} and MdfA C_{In} states (Fig. S3a).

Primary sequence alignment of MdfA and YajR shows 16% identity and 32% consensus (Fig. S1b). Their N-domains show higher homology than their C-domains (which is a general trend observed among MFS proteins [8]), and both conserved motifs A and B are observed in their N-domains. In addition, their loops L6-7 (*i.e.* the inter-domain loop connecting TMs 6 and 7) share recognizable sequence similarity. Thus, these two crystal structures allow us to model the C_{Out} structure of MdfA (and the C_{In} structure of YajR as well) based on domain superposition (Fig. S3b) and provide insights into the conformational changes of MFS antiporters.

Drug resistance assays

In our previous analysis of YajR [6], we hypothesized a functional role of the 3D motif-A in stabilizing the C_{Out} state. In the current study, we constructed more motif-A related mutations in MdfA, in order to test their impact on transport activity. First, mutation of D77 at the essential acidic residue of motif-A, D77A, resulted in loss of drug resistance (Fig. S7b), which is in agreement with previous reports [9]. Furthermore, two glycine residues from either N- or C-domains were mutated, namely G73R and G347D/W. All three mutational variants lost activity, a result consistent with our C_{Out} model (Fig. S7b), in which both $G73^{TM2.5}$ (motif-A) and $G347^{TM11.5}$ (C-domain) are in close contact. In addition, both mutations $S345^{TM11.5}D$ as well as the above-mentioned G347D introduce a self-tethered N-cap for TM11. Such an N-cap possibly competes with the D77 from motif-A for binding to the N-terminal of TM11 (Fig. S7b). According to the motif-A hypothesis,

this competition would disrupt the inter-domain charge-dipole interaction in the putative C_{Out} state [6]. Indeed, we found that these mutant variants lost their drug-resistance activity (Fig. S7b). Mutations R78Q, R81Q, and R198Q were designed to test effects of positive charges from R78, R81 and R198 surrounding the critical acidic residue, D77, of motif-A. The mutant forms R78Q and R198Q lost almost all of their activity, whereas R81Q exhibited slightly reduced activity. In comparison, the mutation E201Q, which is in close proximity to D77, maintained its full activity (Fig. S7b). These results suggest that the positively charged residues surrounding D77 are important for the function of motif-A. Moreover, the amphipathic helix, α6-7, is thought to sense the protonation-induced inward movement of the TM core relative to the membrane, thus regulating motif-A in response to the protonation inside the central cavity [6]. Here, we constructed a mutation in MdfA, termed H-$(GS)_5$, where the entire α6-7 helix (residues 206–215 of the sequence “KELGRDYKLV”) was substituted with $(GS)_5$ repeats. Two additional mutations, H-Q_3 and H-Q_7, were also introduced, in which three and seven hydrophobic residues in the amphipathic helix (residues 205–216 of the sequence “LKELGRDYKLVL”) were mutated to Gln, respectively, resulting in sequences of either “QKEQGRDYKQVL” or “QKEQQRDQKQQQ”. Our data show that all three variants, H-$(GS)_5$, H-Q_3, and H-Q_7 lost their resistance to Cm (Fig. S7e), supporting the idea that the amphipathic helix plays an essential role in regulating the interaction between motif-A and the C-domain [6].

To verify the functional importance of the inter-domain salt-bridge bond $E136^{TM5.5}$-$R336^{TM10.5}$ (predicted to be formed in the putative C_{Out} state (Fig. S3c) and potentially present in nearly all MdfA orthologs (Fig. S1a)), we constructed E136R, R336E, and E136R/R336E mutations. E136R lost nearly all resistance activity, while both R336E and the double mutant E136R/R336E conferred full Cm-resistance (Fig. S7f). Note that it is possible that the side chain of the R336E point mutation may form H-bonds with either R131 (see Supplementary) or E136 in the C_{Out} state. Together, these results support our C_{Out} model (Fig. S3), in

which the E136$^{TM5.5}$-R336$^{TM10.5}$ salt-bridge bond plays an important role in the stabilization of the C_{Out} state.

In addition, a number of conserved Pro residues located in the N-C domain interface in the C_{In} state were mutated to Ala. P243$^{TM7.1}$A maintained full Cm-resistance activity, whereas P38$^{TM1.1}$A was partially functional. Mutations at two conserved positions in the motif-C (*i.e.* the antiporter motif), namely P154$^{TM5.2}$A and P158$^{TM5.1}$A, also resulted in the loss of Cm-resistance (Fig. S7d).

Previous functional studies on MdfA

In a 2012 landmark study on MdfA functions [10], it was observed that low pH inhibits substrate binding (TPP^+ and pyronin$^+$) of detergent-solubilized MdfA (Fig. 1A in Ref. [10]). Substrate binding induces proton release at a stoichiometry of 1:1 (Fig. 2 in Ref. [10]). This is a good example of competition between protonation and substrate binding.

PEGylation is commonly used to investigate solvent accessibility of a given position (mutated to a Cys residue) in a target protein [10]. In MdfA, PEGylation at D34$^{TM1.2}$C is only observed in the presence of TPP^+ (Fig. 5F in Ref. [10]). According to our crystal structure, PEGylated D34C would not fit in the cavity in the C_{In} state because of the large size of PEG (MW: ~5 kDa), but it might occur in the C_{Out} state (Fig. S3b). In contrast, A128$^{TM4.5}$C can be PEGylated only in the absence of TPP^+ (Fig. 1B-D in Ref. [10]). In the C_{In} crystal structure of MdfA, A128 is solvent-exposed and is not directly involved in substrate binding. However, in the C_{Out} state, A128 is located in the putative domain interface. Therefore, we assume that PEGylation at A128C can only occur in the C_{In} state. Similarly, [^{14}C]NEM labeling at Cys residues has been used to assess solvent exposure of a variety positions in MdfA [1]. Together, results of these experiments strongly support that (i) in the absence of substrates, C_{In} is the dominant state of detergent solubilized MdfA; (ii) TPP^+ binding stabilizes the C_{Out} state (*i.e.* $K_d^{Out} < K_d^{In}$); and (iii) binding of either chloramphenicol (Cm) or $EtBr^+$ stabilizes the C_{In} state in the absence of $\Delta\mu(H^+)$,

while their binding favors the C_{Out} state in the presence of $\Delta\mu(H^+)$ [1]. The differential binding energy of substrates may partially contribute to the C_{In}-to-C_{Out} conformational change *in vivo*.

Protonated acidic residues react with *N,N'*-Dicyclohexylcarbodiimide (DCCD; MW: 206 Da, and of two phenol rings) in a hydrophobic environment. The adducts can then be identified using mass-spectroscopy analysis [10]. DCCD reaction and substrate binding were shown to inhibit each other (Figs. 5a, 5c, and 5d in Ref. [10]). In addition, the DCCD reaction was observed at $D34^{TM1.2}$, but not at $E26^{TM1.4}$ (Fig. 6 in Ref. [10]). Based on structure comparisons of both C_{In} and putative C_{Out} states (Figs. 2 and S3), the D34 environment appears to become more hydrophobic in the C_{In} state relative to the C_{Out} state. However, E26 is more solvent-exposed in C_{In} than in the C_{Out} state. Thus, the observed DCCD reaction at D34 is more likely to occur in the C_{In} state than in the C_{Out} state. Should the reaction be able to occur in the C_{Out} state, substrate binding would have no steric conflict on D34 (since PEGylation at D34C has been observed in the TPP^+-stabilized C_{Out} state). Furthermore, the lack of reaction at E26 strongly indicates that E26 is in a deprotonated state in the absence of substrates. Therefore, possible interpretations of the DCCD reaction experiments include (without being mutually exclusive), (i) that stereochemical hindrance between the substrate and DCCD occurs in the C_{In} state; (ii) that substrate induces deprotonation at D34 (*i.e.* the competition hypothesis of antiporters), thus disfavouring the DCCD reaction; and (iii) that TPP^+ binding stabilizes the C_{Out}, whereas the DCCD reaction stabilizes the C_{In} state.

Loss-of-function mutations of E26X can be rescued by a number of secondary point mutations which potentially interrupt the stability of the C_{Out} state [1]. Such rescuing mutations include $V125^{TM4.4}A$, $Y127^{TM4.5}N/H$, $A129^{TM4.5}T$, $A147^{TM5.4}T$, $A150^{TM5.3}V$, and $V335^{TM10.5}M$ in the putative inter-domain interface of the C_{Out} state and both S133F and A191T/V in the vicinity of the so-called PET(S) domain that stabilized the charge-relay triad. These observations are in

agreement with our hypothesis that the interaction between the membrane potential and protonated E26 drives the C_{Out}-to-C_{In} conformational change (Fig. 6). Lowering the energy barrier of the transition state may, to some extent, compensate the loss of the driving force. In addition, C21 of MdfA is shown to become more solvent-exposed in the C_{Out} state (*i.e.* upon TPP^+ binding) than in the C_{In} state (*i.e.* upon Cm binding in the absence of $\Delta\mu(H^+)$) (Fig. 7 of Ref. [1]). On the basis of the crystal structure of MdfA, $C21^{TM1.5}$ is located in the cytosol-membrane interface. Mutations of C21Y/W/F presumably result in stronger hydrophobic interactions between the position 21 and the surrounding lipid bilayer that the WT, thus favouring the C_{In} state. Intriguingly, these mutations can rescue the E26X mutations (Table 1 in Ref. [1]), in agreement with our hypothesis on E26.

Motif-A

The motif-A, “GxxxDRxGRR”, spans from the C-terminal of TM2 to the N-terminal of TM3, and is the most conserved motif in MFS transporters [6, 8]. In MdfA orthologs, this motif is present in the form of “GPlsDrxGRr” (Fig. S1a). Motif-A in the YajR/3WDO crystal structure has been shown to play a critical role in stabilizing the C_{Out} state through an inter-domain charge-dipole interaction [6]. It is likely to play the same functional role in other motif-A containing MFS transporters, including MdfA. For instance, single molecule DEER analysis showed that mutation $D65^{A05}N$ (the superscript indicating the 5^{th} position in the motif-A) in LmrP, another well studied drug-resistance MFS antiporter, results in increased population of the C_{In} state, similar to the effects of lowing pH (when $D65^{A05}$ gets protonated) [11]. Nevertheless, like in many other MFS crystal structures, in the crystal structure of MdfA C_{In} state, the motif-A is not in its functional state. In particular, $D77^{A05}$ is not involved in the charge-dipole interaction with the N-terminal end of TM11; instead, it is surrounded by basic residues from both the motif-A and the inter-domain loop (L6-7), including R78, R81, and R198. However, a motif-A associated charge-relay triad, $D77^{A05}$-$R81^{A09}$-$E132^{TM4.5}$, is well formed in the MdfA crystal structure (Fig. S3c). This

triad is postulated to regulate the inter-domain charge-dipole interaction [6]. According to a web-based OPENSEQ co-evolution analysis [12], the R81-E132 pair is among the position-pairs of the highest possibility to have co-evolved (scaled score = 3.5). Furthermore, on the basis of homologous modelling, we constructed a C_{Out} state model of MdfA, which shed lights on its motif-A functions (Fig. S3c). Mutation $D77^{A05}A$ of ecMdfA abolished drug resistance (Fig. S7b) [9], and E132A shows impaired transport activity [9]. Here, we showed that mutations of basic residues surrounding D77 result in loss of function to varied extents (Fig. S7b), consistent with their roles in neutralizing D77 in the C_{In} state. Other components of the "3D motif-A" include the conserved $G73^{TM2.1/A01}$ and $G347^{TM11.1}$, which ensure close contact between TMs 2 and 11 in the C_{Out} state (Fig. S3c). Mutations at either of these two positions resulted in complete loss of Cm-resistance (Fig. S7b). Moreover, either elimination of helix α6-7 or change of its amphipathic property abolished drug resistance conferred by MdfA (Fig. S7e). Together, these data and numerous previously reported studies on several MFS proteins [6, 9, 13] support the "motif-A hypothesis" [6]. According to this hypothesis, a protonation-induced inward movement of the TM core relatively to the membrane is sensed by the amphipathic helix, α6-7. In response to the inward movement, this helix induces a conformational change in the L6-7 loop. Through the charge-relay triad, the positive charge(s) of the loop weakens the inter-domain, charge-dipole interaction between $D77^{A05}$ and TM11, thus destabilizing the C_{Out} state of MdfA.

References in supplementary material

1 Adler J, Bibi E. Determinants of substrate recognition by the Escherichia coli multidrug transporter MdfA identified on both sides of the membrane. *J Biol Chem* 2004; **279**:8957-8965.

2 Zhao Y, Terry D, Shi L, Weinstein H, Blanchard SC, Javitch JA. Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature* 2010; **465**:188-193.
3 Qin F. Restoration of single-channel currents using the segmental k-means method based on hidden Markov modeling. *Biophysical journal* 2004; **86**:1488-1501.
4 Cummings MD, Farnum MA, Nelen MI. Universal screening methods and applications of ThermoFluor. *J Biomol Screen* 2006; **11**:854-863.
5 Alexandrov AI, Mileni M, Chien EY, Hanson MA, Stevens RC. Microscale fluorescent thermal stability assay for membrane proteins. *Structure* 2008; **16**:351-359.
6 Jiang D, Zhao Y, Wang X *et al.* Structure of the YajR transporter suggests a transport mechanism based on the conserved motif A. *Proceedings of the National Academy of Sciences of the United States of America* 2013; **110**:14664-14669.
7 Sun L, Zeng X, Yan C *et al.* Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* 2012; **490**:361-366.
8 Paulsen IT, Brown MH, Skurray RA. Proton-dependent multidrug efflux systems. *Microbiological reviews* 1996; **60**:575-608.
9 Sigal N, Molshanski-Mor S, Bibi E. No single irreplaceable acidic residues in the Escherichia coli secondary multidrug transporter MdfA. *J Bacteriol* 2006; **188**:5635-5639.
10 Fluman N, Ryan CM, Whitelegge JP, Bibi E. Dissection of Mechanistic Principles of a Secondary Multidrug Efflux Protein. *Molecular cell* 2012.
11 Masureel M, Martens C, Stein RA *et al.* Protonation drives the conformational switch in the multidrug transporter LmrP. *Nature chemical biology* 2013.
12 Ovchinnikov S, Kamisetty H, Baker D. Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *eLife* 2014; **3**:e02030.
13 Zhao Y, Mao G, Liu M, Zhang L, Wang X, Zhang Xuejun C. Crystal Structure of the E. coli Peptide Transporter YbgH. *Structure* 2014; **22**:1-9.
14 Larkin MA, Blackshields G, Brown NP *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; **23**:2947-2948.

15 Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 2009; **25**:1189-1191.
16 Schuster-Bockler B, Schultz J, Rahmann S. HMM Logos for visualization of protein families. *BMC Bioinformatics* 2004; **5**:7.
17 Fluman N, Bibi E. Bacterial multidrug transport through the lens of the major facilitator superfamily. *Biochimica et biophysica acta* 2009; **1794**:738-747.