Conformational variation of the translocon enhancing chaperone SecDF

Kazuhiro Mio[1†], Tomoya Tsukazaki[2,3], Hiroyuki Mori[4], Masaaki Kawata[1], Toshio Moriya[1], Yoshikazu Sasaki[5], Ryuichiro Ishitani[6], Koreaki Ito[7], Osamu Nureki[6†], Chikara Sato[1†]

[1]BRI and ITRI, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8568, Japan, [2]Graduate School of Biological Science, Nara Institute of Science and Technology, Nara 630-0192, Japan, [3]JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan, [4]Institute for Virus Research, Kyoto University, Sakyo-ku, Kyoto 606-8507, Japan, [5]Advanced Technology Division, JEOL Ltd., Akishima, Tokyo 196-0022, Japan, [6]Department of Biophysics and Biochemistry, Graduate School of Science, The University of Tokyo, 2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan, [7]Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto 603-8555

†To whom correspondence should be addressed.

E-mail: kazu.mio@aist.go.jp, Tel: +81-3-3599-8104, fax: +81-3-3599-8099

nureki@ims.u-tokyo.ac.jp, Tel: +81-3-5841-4394, fax: +81-3-5841-8057

ti-sato@aist.go.jp, Tel: +81-29-861-5562, fax: +81-29-861-6478

Abstract

The Sec translocon facilitates transportation of newly synthesized polypeptides from the cytoplasm to the lumen/periplasm across the phospholipid membrane. Although the polypeptide-conducting machinery is formed by the SecYEG-SecA complex in bacteria, its transportation efficiency is markedly enhanced by SecDF. A previous study suggested that SecDF assumes at least two conformations differing by a 120° rotation in the spatial orientation of the P1 head subdomain to the rigid base, and that the conformational dynamics plays a critical role in polypeptide translocation. Here we addressed this hypothesis by analyzing the 3D structure of SecDF using electron tomography and single particle reconstruction. Reconstruction of wt SecDF showed two major conformations; one resembles the crystal structure of full-length SecDF (F-form structure), while the other is similar to the hypothetical structural variant based on the crystal structure of the isolated P1 domain (I-form structure). The transmembrane domain of the I-form structure has a scissor like cleft open to the periplasmic side. We also report the structure of a double cysteine mutant designed to constrain SecDF to the I-form. This reconstruction has a protrusion at the periplasmic end that nicely fits the orientation of P1 in the I-from. These results provide firm evidence for the occurrence of the I-form in solution and support the proposed F- to I-transition of wt SecDF during polypeptide translocation.

Now 218 words; should be 150 – 250 words

Keywords

Conformational change, EM, SecDF, single particle analysis, translocon

Abbreviations

3D	Three dimensional
CTF	Contrast transfer function
DDM	*n*-dodecyl-β-D-maltoside
DF	Dark field
EM	Electron microscopy
FSC	Fourier shell correlation
MSA-HAC	Multivariate statistical analysis in combination with hierarchical ascendant classification
RND	Resistance-nodulation-cell division
SIRT	Simultaneous iterative reconstruction technique
SPA	Single particle analysis
STEM	Scanning transmission electron microscopy
TEM	Transmission electron microscopy
TM	Transmembrane
TSecDF	*Thermus thermophilus* SecDF
wt	wild type

Introduction

Translocation of newly synthesized secretion proteins across the membrane and integration of membrane proteins into the plasma membrane are essential for all living organisms. These processes are carried out by a specific protein complex called the translocon machinery. Bacterial protein translocation depends on the action of at least three components; the SecYEG complex constructs a peptide-conducting pore, the SecA ATPase generates the driving force for peptide translocation through the pore, and SecDF accelerates the overall translocation reaction [1].

Although the translocation pore is formed by SecYEG complex, and peptide translocation can be driven by the SecA ATPase, the auxilliary membrane protein SecDF is indispensable for the efficient transport of proteins. Genetic depletion of SecDF causes severe growth retardation and protein export defects [2] in a cold sensitive manner [3, 4]. Thus, the bacterial SecDF complex is a plausible target for drug design aimed at preventing growth of pathogenic bacteria.

SecDF belongs to the resistance-nodulation-cell division (RND) superfamily and shares several structural features with other members of the family, such as AcrB that transports drugs and other hydrophobic compounds, CzcA that transports heavy metals, and NolGHI that transports lipooligosaccharides [5]. Most RND proteins have a long polypeptide chain (700-1300 amino acid residues long) with a transmembrane (TM) segment at the N-terminus followed by a large periplasmic domain, six additional TM segments, a second large periplasmic domain, and five C-terminal TM segments. The amino acid sequence of the first half of RND proteins is homologous to the second half. The RND protein homologue from the archaebacterium *Methanococcus jannaschii* is predicted to have six TM domains, and is postulated to function as a homo- or

heterodimer. Thus, the protein is probably the result of a duplication that occurred prior to divergence of the family members. SecDF from *E. coli* consists of two polypeptide chains, while that from *Bacillus subtilis* is a single chain [5]. SecDF of *T. thermophilus* HB8 (TSecDF) exists as a single polypeptide chain of 735 amino acid residues.

The structure of full-length TSecDF has been determined at 3.3 Å resolution [6]. The complex has a pseudo-symmetric, 12-helix TM domain and six periplasmic regions, of which the first and fourth, P1 ($TSecDF_{31-263}$) and P4 ($TSecDF_{469-559}$), are characteristically large (Fig 1a). In P1 subdomain, $TSecDF_{36-111}$ and $TSecDF_{251-263}$, adjacent to the first and the second TM segments, respectively, construct a rigid base from which the P1 head domain protrudes. The head and base subdomains are connected by two linkers that might act as hinges. The segment of the P1 domain extending from residue number 36 to 263 has also been crystallized [6, 7]. Notably, the conformation of P1 in this 2.6 Å structure differs from that in full-length wild type (wt) TSecDF. In the full-length protein most of the P1 head domain leans out away from the central axis of the TM segments (Fig.1, b, left). This conformation is termed the F-form of TSecDF. When the structure of the isolated P1 domain is docked to the rest of the full-length protein, its head domain is rotated by 120° relative to the F-form (Fig.1, b, right). This hypothetical structure is termed the I-form of TSecDF [6]. The result evokes the possibility that TSecDF can exist in at least two different states with distinct P1 configurations. Because the P1 domain can bind an unfolded protein [6], such conformational variance probably has a direct role in the acceleration of polypeptide translocation by TSecDF.

To examine whether the P1 domain of the TSecDF undergoes a conformational change during the polypeptide translocation cycle, we have analyzed the structure of full-length wt TSecDF by electron microscopy (EM) techniques. Further, we introduced a double cysteine mutation designed to constrain the P1 domain to the I-form, and analyzed the structure. The results show that TSecDF indeed exists in the F- and I-forms in solution, supporting the idea that this conformational flexibility of the P1 domain is important for the functioning of this translocation-enhancing Sec factor.

Materials and Methods

Purification of TSecDF

Purification of TSecDF was carried out as described previously [8]. Plasmid pTT206 that contains the *secDF* gene (gene ID 3168575) from *T. thermophilus* HB8 and a His_6 sequence at its C-terminus was introduced into *E. coli* strain AD202 [9] which additionally harboured pSTD343 ($lacI^q$) [10]. Total membranes from the cells [11] were solubilized with 20 m*M* Tris-HCl buffer pH 8.1 containing 2%(*w*/*v*) *n*-dodecyl-β-D-maltoside (DDM), 5%(*v*/*v*) glycerol, 300 m*M* NaCl, 20 m*M* imidazole pH 7.0 and 0.1 m*M* Pefabloc (Merck). After the removal of insoluble material by ultracentrifugation, the supernatant was applied to a Ni-NTA column equilibrated with buffer *A* [20 m*M* Tris-HCl pH 8.1, 0.1%(*w*/*v*) DDM, 5%(*v*/*v*) glycerol, 300 m*M* NaCl and 0.1 m*M* Pefabloc]. The sample was eluted with a 20-300 m*M* imidazole gradient in buffer *A*. Sample was further purified by Hi-Trap SP (GE healthcare) cation-exchange column chromatography, followed by Superose 6 (GE healthcare) size-exclusion column chromatography. Substitutions of leucine 106 to cysteine and leucine 243 to cysteine were conducted by the site-directed mutagenesis. The expression and purification of double-cys SecDF protein was performed similarly.

Electron microscopy

To exclude aggregates or deformed molecules, samples were further analyzed by Superdex 200 size-exclusion chromatography with buffer *A* just prior to the EM analysis. The wt TSecDF protein or a double cysteine mutant (~50 μg/ml) was adsorbed to a glow discharged thin carbon film supported by a copper 400-mesh grid, and

negatively stained with 2% uranyl acetate. Micrographs were recorded using a JEOL 100CX transmission electron microscope (TEM) at 100-kV acceleration voltage, 55,100x magnification, and an electron dose of 30 electrons/Å^2. Kodak SO-163 films and Kodak Developer D-19 were employed. Negatives were digitized at 1.73 Å / pixel at the specimen level.

Particles for the F- and I-forms of TSecDF were separately selected from the raw images when their correlation coefficient with a projection of the corresponding average tomographic volume (see below) was >0.325. The classified raw images were subframed into 120 ×120 pixels, and corrected for the contrast transfer function (CTF). For each reconstruction, the Euler angles of the constituting raw images were determined by the projection matching method, using 91 projections at 15^o increments from the tomographic volume as the initial three dimensional (3D) references. The raw images were classified by their Euler angles, averaged and used to reconstruct the 3D structures by the Simultaneous Iterative Reconstruction Technique (SIRT) [12, 13]. The next set of Euler angles was determined with the reconstructed 3D volume by the projection matching method, and the cycles from 3D-reconstruction to projection matching were iterated until convergence. The FSC was used to evaluate the resolution, with a threshold of 0.5 [14].

STEM tomography

Negatively stained microscopy grids carrying TSecDF were prepared as detailed above. Scanning transmission electron microscopy (STEM) dark-field (DF) tomograms were recorded using a JEM-2100F at an acceleration voltage of 200 kV using a STEM-DFI detector (JEOL) at 3,000,000x, corresponding to a pixel size of 0.98 Å at the

specimen level. Acquisition of a tomographic data set of each particle centered in the view was controlled by the TEMography software (JEOL System Tech.); the grid was tilted with an increment angle of 2 degrees, between –60° and +60° tilt angles on a single tilt axis (61 exposures per dataset). 3D reconstruction was carried out by the SIRT. Each 3D volume was aligned rotationally and translationally using a 3D alignment algorithm[12]. This searched the full range of Euler angles and 10% range for translational shift along each of the three axes (x, y, z). Multivariate statistical analysis followed by the hierarchical ascendant classification (MSA-HAC) algorithm [15] was applied to classify the STEM tomography. Each (120 x 120 x 120) pixel tomogram was fist filtered using a Gaussian Filter and binarized. After this preprocessing, MSA reduced the dimensions of each tomogram to 24. HAC was performed using the 24 dimensional coordinates of each tomogram, and separated the tomograms into two groups. From 38 reconstructed particles, 13 and 11 particles were classified as F- and I-forms, respectively. In each group, the particles were oriented in various ways on the carbon surface, so that the overlapped volume data filled in the missing wedge of each particle. The aligned volumes were averaged with an equal weight. To assess the resolution of the final 3D map, two independent average volumes within a group were compared using the FSC function [14].

Results

Analysis of full length wt TSecDF by electron tomography

To see whether TSecDF adopts multiple conformations in solution, we first observed the structure using electron tomography. EM grids carrying purified TSecDF were negatively stained and imaged by DF STEM. Single tilt tomograms were recorded with an increment angle of 2 degrees, between –60 and +60 degrees (61 exposures per dataset) (Materials and Methods; Fig. 2a). In all, 38 usable TSecDF datasets were obtained and reconstructed using the SIRT. From the 38 reconstructed particles, 24 were clearly classified into two groups by MSA-HAC [15]: 13 and 11 structures into the F- and I-form, respectively, as indicated by the red boundary (Figs 2b and c). Views of the particles are shown in Figs. 2 d and e. The other structures were ambiguous and were omitted from the subsequent averaging step. The particles in a group were oriented in different ways on the carbon surface with the result that the volume data overlapped to fill in the missing wedge present for each particle. The aligned volumes were averaged with an equal weight. The axe-shaped reconstruction resulting from 13 structures (Fig. 2f) is probably the F-form of TSecDF; compare the X-ray structure shown in Fig 1b, left. If this is correct, the ‚blade' marked with an asterisk in Fig. 2f corresponds to the head of the P1 domain. The occurrence of a second conformation (Fig. 2g) indicates that the spatial relationship between the head and base subdomains of P1 is highly flexible, as speculated. This second structure is considered to be the I-form of TSecDF and seems to have one main body with two protrusions (Fig. 2g). According to the Fourier shell correlation (FSC) > 0.5 criterion [14], the resolutions of the two average reconstructed 3D volumes are 31.8 Å for the F-form and 31.4 Å for the I-form.

Single particle analysis of full length wt TSecDF

We next analyzed the structure of full length wt TSecDF by single particle analysis (SPA) to obtain higher resolution data. Due to the relatively low molecular mass of TSecDF (MW 79.7 kDa), negatively stained particles were examined instead of an ice-embedded specimen. A total of 4,910 projection images were selected from digitized EM micrographs and classified by cross-correlation to the 3D volumes obtained by tomography. Projections with a cross-correlation coefficient > 0.325 to one or other of these tomographic volumes (Fig. 2f, g) were used for the reconstructions. In all 2,966 images were classified to the F-form structure and 1,944 images to the I-form structure.

The 3D structures of these particles were reconstructed by SPA without imposing symmetry. Initially, almost all particles were used. However, a highly swollen volume emerged due to associated lipid and detergent; 50% of the TSecDF sequence is located in the membrane. The level of swelling varied depending on how the EM grid was washed and stained after the sample had been loaded. Therefore, in the later analysis swollen particles were strictly discarded.

The final reconstruction of the F form by SPA again displays the axe-shaped structure (Fig. 3a, b). When the atomic coordinates of TSecDF (PDB: 3AQP) were docked into the density, the overall agreement was good (Fig. 3c). However, the P1 domain would fit better if it was tilted towards the membrane by a further 15°. This supports the idea that the P1 domain is flexible. The 2D averages that did not fit to the reconstruction were omitted from the calculation and the particle images of the minor population seem to be lost during the calculation; they are interpreted to be minor conformational variants of SecDF. Intriguingly, unlike the tightly packed F-form the TM

domain of the I-form has a crack at the interface between the SecD and SecF halves (Fig. 3d, asterisk). The crack opens to the periplasm, suggesting a scissor-like movement of the SecD and SecF TM helices coupled with the upright conformational change in the P1 domain. Structural correspondence was shown by agreement between the raw TSecDF images and projections from the 3D reconstructions (Fig.3b, e). The FSC between the independent reconstructions calculated from the two half-sets of the SPA data for the two forms indicates that the resolution limit is 14.1 Å for the F-form (Fig. 3g) and 14.8 Å for the I-form by the FSC > 0.5 criterion (Fig. 3h). This is much better than the resolution of the 3D volumes obtained by tomography (31.8Å for the F-form, and 31.4 Å for the I-form).

Cross-correlation of 1000 randomly picked particles with the 3D volumes obtained by tomography indicated the prevalence of the two TSecDF forms in solution. In all 43% of TSecDF particles classified as either the F- or the I-form when the cross-correlation coefficient was > 0.325, and this ratio increased up to 71% when a threshold of 0.300 was adopted (Table I). Considering that a correlation coefficient of 0.300 is high enough to detect structural similarity, these data imply that the F- and I-forms are major TSecDF conformations in solution.

Structure of a TSecDF double cystein mutant

As EM tomography showed structural variance in TSecDF, especially in the particle group classified as the I-form (Fig. 2c), we used disulfide bond formation to try and fix the protein in a specific conformation. Previous experiments showed that disulfide bond formation by double-cysteine mutants of *E. coli* SecD increased with time, probably giving rise to the I-form, and that this markedly impaired protein export [6]. The

structure of a double-cys (Leu106Cys/Leu243Cys) TSecDF mutant was determined; the new cysteine residues are closest in the I-form of TSecDF (Fig. 4a). Most part of the mutant protein ran slightly faster than the wild type protein when SDS-PAGE was performed under non-reducing conditions (Fig. 4b). As the change in molecular weight caused by the amino acid substitution is negligable, this modulation is considered to be caused by a conformational change in the TSecDF molecules resulting from the formation of a disulfide bond between the new cysteins. From the density of the bands, the disulfide fixation occurs in more than 90% of the protein. The mutant protein was prepared for negative stain EM and imaged for SPA. Initially, 11,194 particle projections were manually selected from micrographs recorded by negative stain TEM. A stricter selection was made again to avoid over swelling of the 3D reconstruction (see above). The final 3D structure was reconstructed from the 4,555 particles.

The reconstructed 3D structure has a large TM domain at the bottom and a protrusion at the periplasmic side (Fig. 4c). Docking the wt TSecDF structure into the volume showed the conformation of the P1 domain in the I-form to fit with the uppermost density best, the F-form has no corresponding density (Fig. 4d, right). The FSC > 0.5 criterion shows a resolution limit at 20.8 Å (Fig. 4e).

Discussion

In the 3.3 Å crystal structure of full-length wt TSecDF the head and base of the P1 domain are connected by two linkers that were proposed to form a flexible hinge [6]. Further, docking the crystal structure of the P1 domain alone into this F-form TSecDF structure indicated that a second conformation may be possible. This hypothetical structure was termed the I-form. In the I-form the head of the P1 domain is rotated by 120° relative to its position in the F-form. These studies also suggested that the crosslinks between the P1 head and base domains into predicted I-form can be formed by time-dependent manner [6], suggesting F-I transition in solution. However, actual conformational dynamics of the P1 domain was not directly observed. To address this question, we purified wt TSecDF and investigated its structure by electron tomography and negative staining transmission electron microscopy (TEM). The data from both revealed the presence of two distinct TSecDF structures in solution (Figs. 2 and 3) : one similar to the F-form and the other to the proposed I-form (Fig. 1). An independent assay where 1000 TSecDF particles were randomly selected (Table 1), suggests that together the F- and I-forms are the major conformations of TSecDF in solution, even when the oxidative conditions reported to be a requirement for the F-form [6] are not met. The shape of the 3D volumes of wt TSecDF obtained for the F-form by electron tomography (Fig. 2) and at higher resolution by TEM and SPA (Fig. 3) matched the crystal structure (Fig. 1) well, while the structure of the double-cys TSecDF mutant (Fig. 4) was similar to the proposed I-form (Figs. 1, 2 and 3).

SecDF belongs to the RND family of multidrug exporters [5]. Characterized members of the RND superfamily all probably catalyze substrate efflux via a proton

antiport mechanism, and some aspects of the export mechanisms must be common to them all. One transport system of the RND family is well documented, i.e., the *E. coli* multidrug efflux transporter AcrB [16, 17]. Three AcrB protomers are organized as a homotrimer. Each protomer consists of a 50 Å thick TM domain and a 70 Å headpiece, protruding from the external membrane surface. Crystallographic studies of the asymmetric trimer of AcrB suggest that each protomer goes through a cycle of conformational changes during drug export [18, 19]. Site-directed disulfide cross-linking further demonstrated that the flexibility of the periplasmic domain is essential for drug efflux [18]. Peptide transport by the Sec translocon is coupled to proton transport, as in AcrB and other RND transporters. Furthermore, a disulfide crosslink in the periplasmic P1 domain of TSecDF causes severe transport deficiency [6]. In the present work we confirmed the presence of two different conformations of TSecDF in solution. Analysis of time-lapse movement of this domain is now required.

In contrast to the spatial translocation of the P1 domain, the conformational changes in other parts of the molecule is not yet clear. The I-form reconstructed from images of wt TSecDF has a crack open to the periplasm (Fig. 3) that is not present in the F-form, suggesting that the SecD and SecF TM helices move in a scissor-like manner. As is known from the X-ray structure of TSecDF [6], the periplasmic side of the TM interface involves an array of hydrophobic interactions that could preclude the potential TM proton flow. Therefore, the crack in the I form may facilitate easier access of protons to the TM region. P1 movement at the periplasm and a spatial translocation of the intramembrane helices may be coupled with the proton antiport. X-ray structures of the bacterial ABC lipid flippase, MsbA, demonstrate such a large conformational changes in the membrane during substrate transport [20]. Although TSecDF is not an independent

transporter, and cannot transport peptide by itself, the possibility of a large conformational change in the membrane synchronized to the peptide transport cycles cannot be excluded.

Double-cys mutation was introduced to the TSecDF expecting that the structure to be fixed similar to that of the native I-form. We showed in this study that their overall configurations are similar, but the TM helices of double-cys TSecDF have a tighter packing than those of the wt I-form (Fig. 4). However, the attachment of a large volume of lipids and detergent to the hydrophobic TM domains, and the relatively low resolution EM analysis excludes more precise discussion about the structural difference. As demonstrated here for TSecDF, the dynamic conformational change of a protein can be analyzed by a combination of EM and X-ray crystallography. This methodology should be more widely applied to various kinds of proteins to gain deeper insights into the dynamism that they undergo.

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas, Structural basis of cell-signaling complexes mediating signal perception, transduction and responses, by grants from PREST and CREST, from the Ministry of Economy, Trade and Industry, and from the Ministry of Education, Culture, Sports, Science, and Technology.

	F-form	I-form	Others	Total
CCF > 0.325	195	231	574	1,000
	(19.5)%	(23.1)%	(57.4)%	(100)%
CCF > 0.300	339	374	287	1,000
	(33.9)%	(37.4)%	(28.7)%	(100)%

Table I. Classification of 1000 randomly picked particles to the F-form and I-form of SecDF according to the correlation coefficient.

References

1. du Plessis DJ, Nouwen N, and Driessen AJ (2011) The Sec translocase. Biochim Biophys Acta 1808:851-865
2. Pogliano JA, and Beckwith J (1994) SecD and SecF facilitate protein export in *Escherichia coli*. EMBO J 13:554-561
3. Bolhuis A, Broekhuizen CP, Sorokin A, van Roosmalen ML, Venema G, Bron S, Quax WJ, and van Dijl JM (1998) SecDF of *Bacillus subtilis*, a molecular Siamese twin required for the efficient secretion of proteins. J Biol Chem 273:21217-21224
4. Pogliano KJ, and Beckwith J (1993) The Cs sec mutants of *Escherichia coli* reflect the cold sensitivity of protein export itself. Genetics 133:763-773
5. Tseng TT, Gratwick KS, Kollman J, Park D, Nies DH, Goffeau A, and Saier M H, Jr. (1999) The RND permease superfamily: an ancient, ubiquitous and diverse family that includes human disease and development proteins. J Mol Microbiol Biotechnol 1:107-125
6. Tsukazaki T, Mori H, Echizen Y, Ishitani R, Fukai S, Tanaka T, Perederina A, Vassylyev DG, Kohno T, Maturana AD, Ito K, and Nureki O (2011) Structure and function of a membrane component SecDF that enhances protein export. Nature 474:235-238
7. Echizen Y, Tsukazaki T, Dohmae N, Ishitani R, and Nureki O (2011) Crystallization and preliminary X-ray diffraction of the first periplasmic domain of SecDF, a translocon-associated membrane protein, from *Thermus thermophilus*. Acta Crystallogr Sect F Struct Biol Cryst Commun 67:1367-1370
8. Tsukazaki T, Mori H, Fukai S, Numata T, Perederina A, Adachi H, Matsumura

H, Takano K, Murakami S, Inoue T, Mori Y, Sasaki T, Vassylyev DG, Nureki O, and Ito K (2006) Purification, crystallization and preliminary X-ray diffraction of SecDF, a translocon-associated membrane protein, from *Thermus thermophilus*. Acta Crystallogr Sect F Struct Biol Cryst Commun 62:376-380

9. Akiyama Y, and Ito K (1990) SecY protein, a membrane-embedded secretion factor of *E. coli*, is cleaved by the ompT protease in vitro. Biochem Biophys Res Commun 167:711-715
10. Mori H, Akiyama Y, and Ito K (2003) A SecE mutation that modulates SecY-SecE translocase assembly, identified as a specific suppressor of SecY defects. J Bacteriol 185:948-956
11. Matsuo E, Mori H, Shimoike T, and Ito K (1998) Syd, a SecY-interacting protein, excludes SecA from the SecYE complex with an altered SecY24 subunit. J Biol Chem 273:18835-18840
12. Penczek P, Radermacher M, and Frank J (1992) Three-dimensional reconstruction of single particles embedded in ice. Ultramicroscopy 40:33-53
13. Scheres SH, Valle M, Nunez R, Sorzano CO, Marabini R, Herman GT, and Carazo JM (2005) Maximum-likelihood multi-reference refinement for electron microscopy images. J Mol Biol 348:139-149
14. Harauz G, and van Heel M (1986) Exact filters for general geometry 3-dimensional reconstruction. Optik 73:146-156
15. Seo J, Shneiderman B, Yamashita E, Matsumoto T, and Yamaguchi A (2006) Knowledge discovery in high-dimensional data: case studies and a user survey for the rank-by-feature framework. IEEE Trans Vis Comput Graph 12:311-322
16. Murakami S, Nakashima R, Yamashita E, Matsumoto T, and Yamaguchi A

(2006) Crystal structures of a multidrug transporter reveal a functionally rotating mechanism. Nature 443:173-179

17. Murakami S, Nakashima R, Yamashita E, and Yamaguchi A (2002) Crystal structure of bacterial multidrug efflux transporter AcrB. Nature 419: 587-593

18. Takatsuka Y, and Nikaido H (2007) Site-directed disulfide cross-linking shows that cleft flexibility in the periplasmic domain is needed for the multidrug efflux pump AcrB of *Escherichia coli*. J Bacteriol 189:8677-8684

19. Takatsuka Y, Chen C, and Nikaido H (2010) Mechanism of recognition of compounds of diverse structures by the multidrug efflux pump AcrB of *Escherichia coli*. Proc Natl Acad Sci U S A 107:6559-6565

20. Ward A, Reyes C L, Yu J, Roth C B, and Chang G (2007) Flexibility in the ABC transporter MsbA: Alternating access with a twist. Proc Natl Acad Sci U S A 104:19005-19010

Figure Legends

Fig. 1 Structure of TSecDF. a, The protein backbones of full-length TSecDF (PDB: 3AQP) viewed from the side are shown as wire diagrams. The predicted position of the 30 Å wide membrane is indicated. **b,** (left) Crystal structure of the full-length TSecDF F-form, and (right) the predicted structure of the full-length TSecDF I-form. Rigid-body rotation of the P1 head domain of the F-form by 120° towards the periplasmic space was required to dock the 2.6 Å resolution crystal structure of the isolated P1 domain (PDB: 3AQO) into the structure

Fig. 2 Electron tomography of TSecDF revealing the F- and I-forms. a, TEM images of negatively stained TSecDF. **b,** Hierarchical clustering of the TSecDF tomographic reconstructions. Dendrogram showing the classification of the 24 TSecDF reconstructions by MSA-HAC. Character of each volume was extracted as the coordinate in 24 dimensional feature space for this (Materials and Methods). The reconstructions separated into two groups as indicated by the dotted red line. The resultant classification was used to reconstruct the average 3D models shown in *f* and *g*. **c,** Cluster analysis of the 24 three dimensionally aligned TSecDF volumes: map of the second versus the third factor. The dotted red line delineates the groups indicated in *b*. **d, e,** Gallery showing comparable views of the 3D particle reconstructions obtained from the DF STEM tomography data by the SIRT, and classified as F-form (d; 13 particles) and I-form (e; 11 partilces) by MSA-HAC. Each volume was contoured at an isosurface enclosing 100% of the total volume of TSecDF (79.7 kDa), and viewed from the side. **f, g**, Cluster averages of F-form and I-form TSecDF obtained using the particles shown in *d* and *e*, respectively. Scale bars: 200 Å in a, and 50 Å in d, e, f and g

Fig. 3 Single particle reconstruction of wt TSecDF. a, 3D structure of F-form TSecDF reconstructed from negatively stained wt TSecDF. **b,** A representative array of raw images (upper panels) with the corresponding reprojections from the F-form 3D reconstruction (lower panels). Good consistency was observed. **c,** Superimposition of the atomic coordinates of TSecDF (PDB: 3AQP) on the reconstructed F-form structure. **d,** 3D structure of I-form TSecDF. Note the scissor like cleft at the transmembrane domain of the I-form reconstruction (asterisk). **e,** Raw images with the corresponding reprojections from the I-form reconstruction. **f**, The reconstructed I-form structure was fitted with the segments from PDB 3AQP and 3AQO. **g,** The FSC function between the two independent reconstructions calculated from the two halves of the data. A resolution limit of 14.1 Å was estimated for the F-form, and **h,** 14.8 Å for the I-form by the FSC > 0.5 criterion. The predicted position of the plasmamembrane is indicated by blue lines in a and d; their saparation corresponds to 30 Å. Scale bars represent 50 Å

Fig. 4 Single particle reconstruction of double-cys TSecDF a, Spatial arrangement of the P1 head domain in F-form (orange), and I-form (magenta) TsecDF [6]. Site-directed cysteine substitutions were introduced at Leu106 and Leu243. The distance between the Cß atoms of the residues are 20.2 Å in the F-form and 7.5 Å in the I-form, indicating that intramolecular cysteine crosslinking can constrain the periplasmic site to the I-form like structure. **b**, Non reducing SDS-PAGE; more than 90% of the double-cys TSecDF protein formed intramolecular crosslinks. **c**, Reconstructed 3D volume of double-cys TSecDF from a negatively stained specimen. **d,** (left) Predicted structure of the I-form. The P1 domain of the original F-form composition (PDB: 3AQP) is shown in

transparent orange. (right) Docking of the same structure into the reconstructed double-cys TSecDF. Note the P1 domain of the proposed I-form (magenta) fits well into the upper protrusion of the EM reconstruction, while the P1 of the F-form (orange) has no corresponding density. **d,** A resolution limit of 20.8 Å was estimated by the FSC > 0.5 criterion. Scale bar in c represents 50 Å

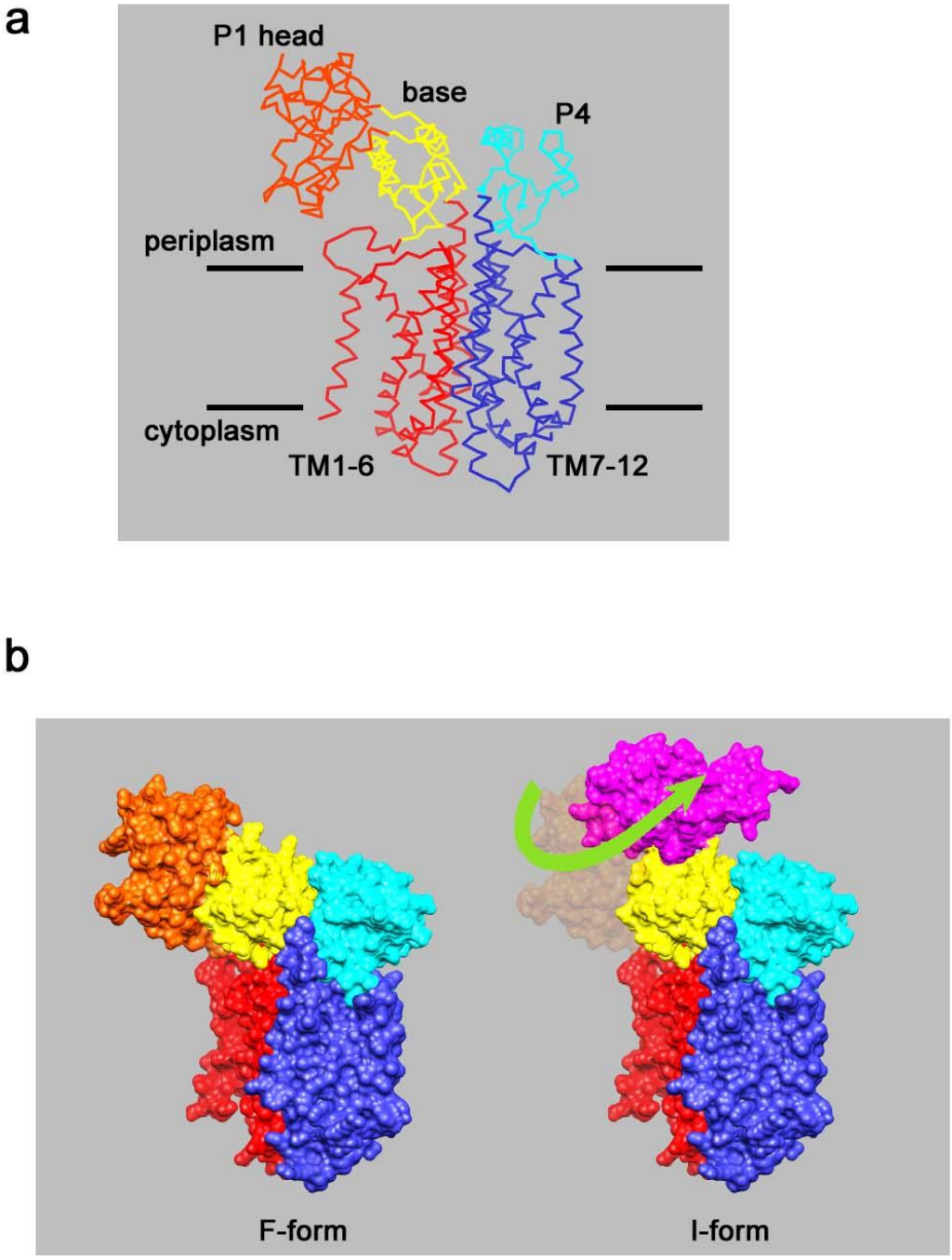
a
P1 head
base
P4
periplasm
cytoplasm
TM1-6
TM7-12
b
F-form
I-form

Figure-1

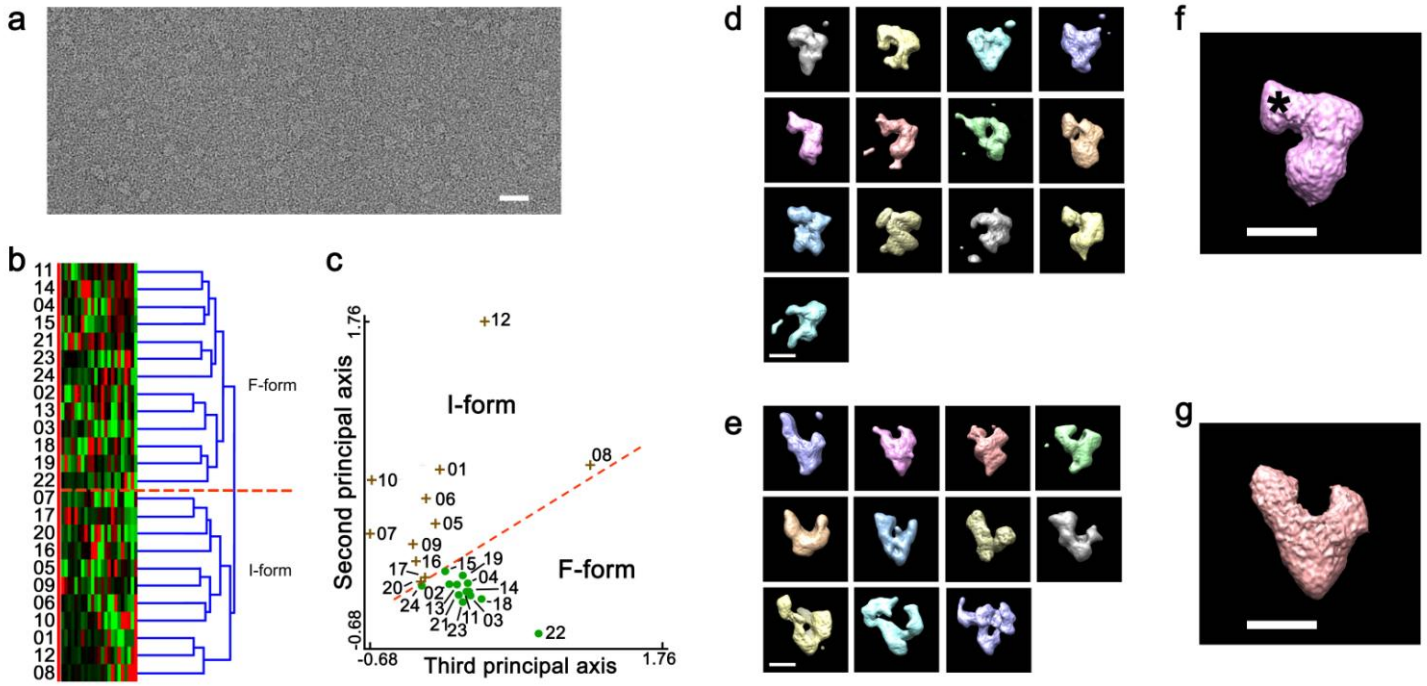
a
b
11
14
04
15
21
23
24
02
13
03
18
19
22
07
17
20
16
05
09
06
10
01
12
08
F-form
I-form
c
Second principal axis
1.76
-0.68
I-form
F-form
+12
+01
+08
+10
+06
+05
+07
+09
+16
+15
+19
17
-04
20
02
-14
24
13
-18
11
03
21
23
22
-0.68
Third principal axis
1.76
d
e
f
g

Figure-2

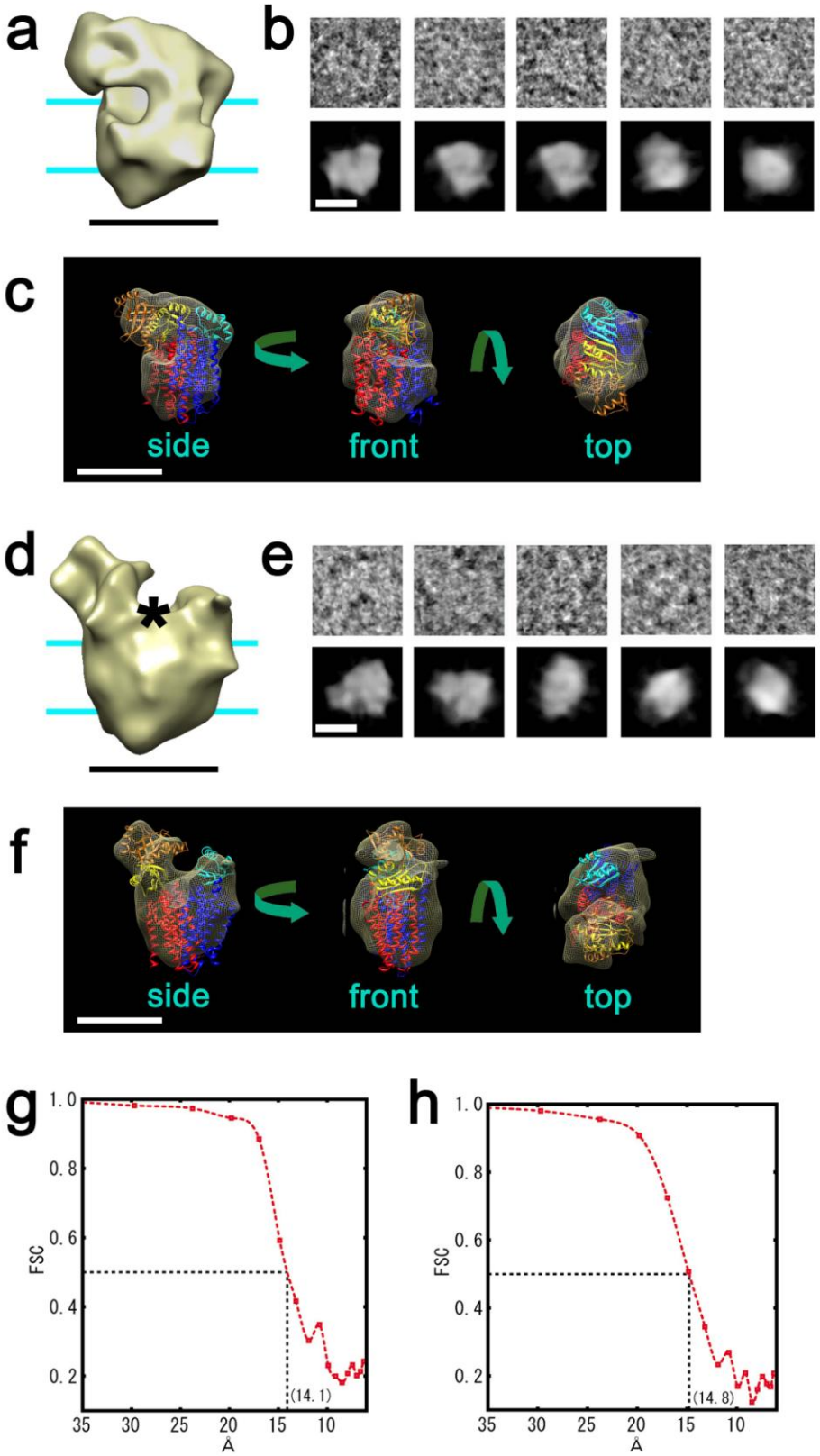
a
b
c
side
front
top
d
e
f
side
front
top
g
h
1.0
0.8
0.6
0.4
0.2
FSC
(14.1)
(14.8)
35
30
25
20
15
10
Å

Figure-3

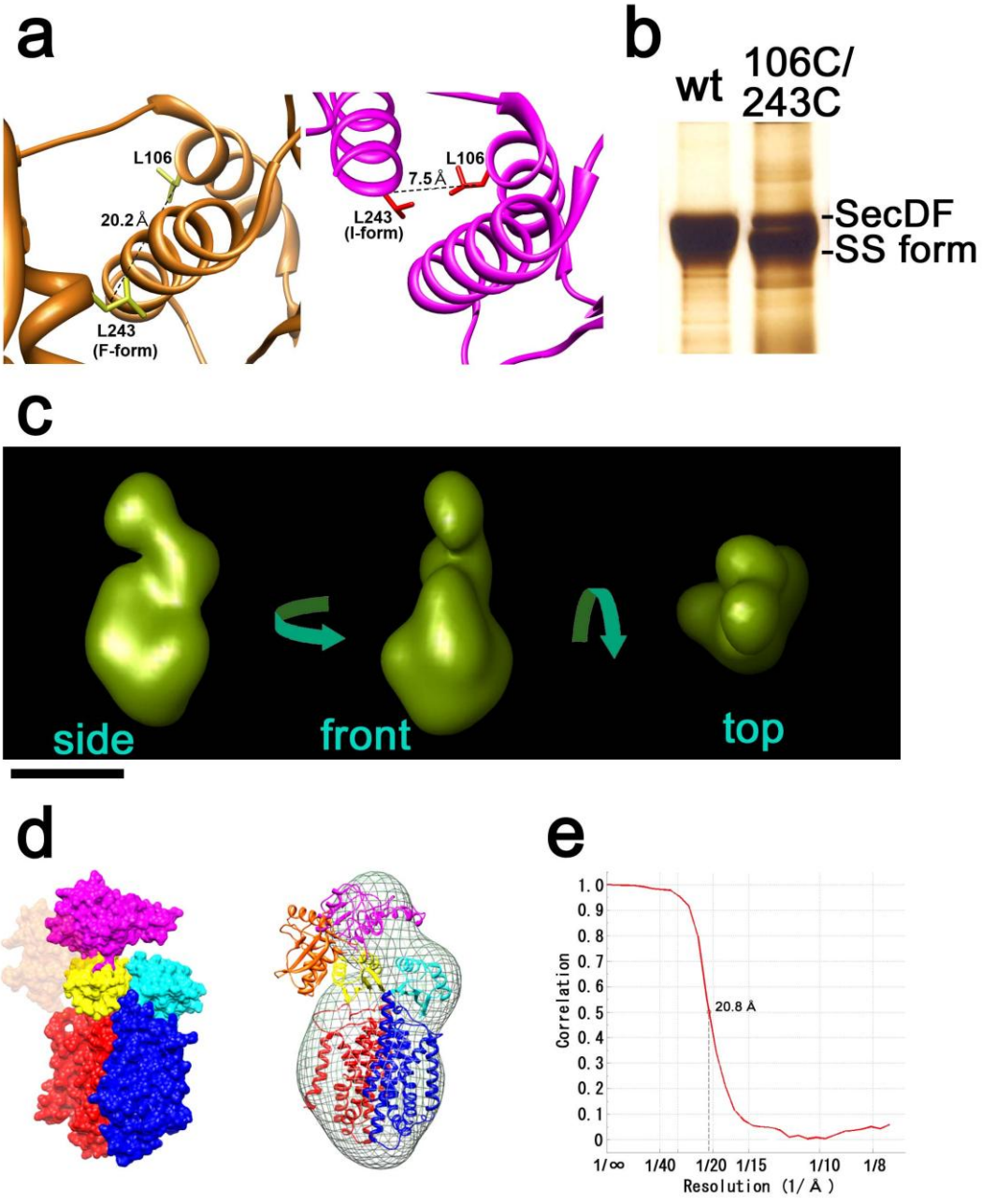
a
L106
20.2Å
L243
(F-form)
L106
7.5Å
L243
(I-form)
b
wt
106C/
243C
-SecDF
-SS form
c
side
front
top
d
e
Correlation
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
20.8 Å
1/∞
1/40
1/20
1/15
1/10
1/8
Resolution (1/Å)

Figure-4