progression is elevated by leptin signaling that is upregulated by obesity and overeating.

The mechanism of LEPR expression during the progression of ACF to colonic tumors is still unknown and further analysis is needed. Our results obtained with the use of colonic epithelium-specific LEPR-deficient mice clearly indicated that LEPR-mediated signaling is more important for the progression of ACF to colonic tumors than for the formation of ACF from normal colonic epithelium. These findings may provide the answers to all previous questions about the role of LEPR-mediated signaling in the development of colonic tumors. In addition, our findings may suggest novel targets for the development of preventive/therapeutic strategies against CRC, based on the inhibition of LEPR-mediated STAT3 signaling.

In conclusion, we clearly demonstrated the local and precise roles of LEPR-mediated signaling in colorectal carcinogenesis. LEPR-mediated signaling, mediated via activation of STAT3 signaling, is important for the progression of ACF to colonic tumors. Our findings may suggest novel targets for the development of preventive/therapeutic strategies against CRC.

Supplementary material

Supplementary Methods and Supplementary Figures 1–4 can be found at http://carcin.oxfordjournals.org/

References

1. Jemal,A. *et al.* (2011) Global cancer statistics. *CA Cancer J. Clin.*, **61**, 69–90.
2. Anderson,W.F. *et al.* (2003) Colorectal carcinoma in black and white race. *Cancer Metastasis Rev.*, **22**, 67–82.
3. Lee,M.Y. *et al.* (2012) The association of diabetes mellitus with liver, colon, lung, and prostate cancer is independent of hypertension, hyperlipidemia, and gout in Taiwanese patients. *Metabolism*, **61**, 242–249.
4. Larsson,S.C. *et al.* (2005) Diabetes mellitus and risk of colorectal cancer: a meta-analysis. *J. Natl Cancer Inst.*, **97**, 1679–1687.
5. Frezza,E.E. *et al.* (2006) Influence of obesity on the risk of developing colon cancer. *Gut*, **55**, 285–291.
6. Giovannucci,E. *et al.* (1997) The role of fat, fatty acids, and total energy intake in the etiology of human colon cancer. *Am. J. Clin. Nutr.*, **66** (6 suppl.), 1564S–1571S.
7. Flegal,K.M. *et al.* (2007) Cause-specific excess deaths associated with underweight, overweight, and obesity. *JAMA*, **298**, 2028–2037.
8. Nam,S.Y. *et al.* (2010) Abdominal visceral adipose tissue predicts risk of colorectal adenoma in both sexes. *Clin. Gastroenterol. Hepatol.*, **8**, 443–450.
9. Frezza,E.E. *et al.* (2006) Influence of obesity on the risk of developing colon cancer. *Gut*, **55**, 285–291.
10. Zhang,Y. *et al.* (1994) Positional cloning of the mouse obese gene and its human homologue. *Nature*, **372**, 425–432.
11. Robert V. Considine, *et al.* (1996) Serum immunoreactive-leptin concentrations in normal-weight and obese humans. *N Engl J Med.*, **334**:292–295.
12. Hausman,G.J. *et al.* (2012) Leptin and reproductive function. *Biochimie*, **94**, 2075–2081.
13. Maffei,M. *et al.* (1995) Leptin levels in human and rodent: measurement of plasma leptin and ob RNA in obese and weight-reduced subjects. *Nat. Med.*, **1**, 1155–1161.
14. Tamakoshi,K. *et al.* (2005) Leptin is associated with an increased female colorectal cancer risk: a nested case-control study in Japan. *Oncology*, **68**, 454–461.
15. Stattin,P. *et al.* (2004) Obesity and colon cancer: does leptin provide a link? *Int. J. Cancer*, **109**, 149–152.
16. Wallace,A.M. *et al.* (1998) Effect of weight loss and the inflammatory response on leptin concentrations in gastrointestinal cancer patients. *Clin. Cancer Res.*, **4**, 2977–2979.
17. Arpaci,F. *et al.* (2002) Low serum leptin level in colon cancer patients without significant weight loss. *Tumori*, **88**, 147–149.
18. Hardwick,J.C. *et al.* (2001) Leptin is a growth factor for colonic epithelial cells. *Gastroenterology*, **121**, 79–90.
19. Aparicio,T. *et al.* (2005) Leptin stimulates the proliferation of human colon cancer cells *in vitro* but does not promote the growth of colon cancer xenografts in nude mice or intestinal tumorigenesis in Apc(Min/+) mice. *Gut*, **54**, 1136–1145.
20. Rouet-Benzineb,P. *et al.* (2004) Leptin counteracts sodium butyrate-induced apoptosis in human colon cancer HT-29 cells via NF-kappaB signaling. *J. Biol. Chem.*, **279**, 16495–16502.
21. Chaudhary,M. *et al.* (2000) Starvation, leptin and epithelial cell proliferation in the gastrointestinal tract of the mouse. *Digestion*, **61**, 223–229.
22. Hirose,Y. *et al.* (2004) Enhancement of development of azoxymethane-induced colonic premalignant lesions in C57BL/KsJ-db/db mice. *Carcinogenesis*, **25**, 821–825.
23. Aparicio,T. *et al.* (2004) Leptin reduces the development of the initial precancerous lesions induced by azoxymethane in the rat colonic mucosa. *Gastroenterology*, **126**, 499–510.
24. Endo,H. *et al.* (2011) Leptin acts as a growth factor for colorectal tumours at stages subsequent to tumour initiation in murine colon carcinogenesis. *Gut*, **60**, 1363–1371.
25. Uchiyama,T. *et al.* (2011) Leptin receptor is involved in STAT3 activation in human colorectal adenoma. *Cancer Sci.*, **102**, 367–372.
26. Uchiyama,T. *et al.* (2011) Role of the long form leptin receptor and of the STAT3 signaling pathway in colorectal cancer progression. *Int. J. Oncol.*, **39**, 935–940.
27. Ray,M.K. *et al.* (2000) The Cre-loxP system: a versatile tool for targeting genes in a cell- and stage-specific manner. *Cell Transplant.*, **9**, 805–815.
28. McMinn,J.E. *et al.* (2004) An allelic series for the leptin receptor gene generated by CRE and FLP recombinase. *Mamm. Genome*, **15**, 677–685.
29. Fujisawa,T. *et al.* (2008) Adiponectin suppresses colorectal carcinogenesis under the high-fat diet condition. *Gut*, **57**, 1531–1538.
30. Tudek,B. *et al.* (1989) Foci of aberrant crypts in the colons of mice and rats exposed to carcinogens associated with foods. *Cancer Res.*, **49**, 1236–1240.
31. Bird,R.P. (1987) Observation and quantification of aberrant crypts in the murine colon treated with a colon carcinogen: preliminary findings. *Cancer Lett.*, **37**, 147–151.
32. Osawa,E. *et al.* (2003) Peroxisome proliferator-activated receptor gamma ligands suppress colon carcinogenesis induced by azoxymethane in mice. *Gastroenterology*, **124**, 361–367.
33. Hosono,K. *et al.* (2010) Metformin suppresses azoxymethane-induced colorectal aberrant crypt foci by activating AMP-activated protein kinase. *Mol. Carcinog.*, **49**, 662–671.
34. Ochiai,M. *et al.* (1996) DNA adduct formation, cell proliferation and aberrant crypt focus formation induced by PhIP in male and female rat colon with relevance to carcinogenesis. *Carcinogenesis*, **17**, 95–98.

Determination of damage-free crystal structure of an X-ray–sensitive protein using an XFEL

Kunio Hirata[1,2,9], Kyoko Shinzawa-Itoh[3,9],
Naomine Yano[2,3], Shuhei Takemura[3], Koji Kato[3,8],
Miki Hatanaka[3], Kazumasa Muramoto[3],
Takako Kawahara[3], Tomitake Tsukihara[2–4],
Eiki Yamashita[4], Kensuke Tono[5], Go Ueno[1],
Takaaki Hikima[1], Hironori Murakami[1],
Yuichi Inubushi[1], Makina Yabashi[1], Tetsuya Ishikawa[1],
Masaki Yamamoto[1], Takashi Ogura[6], Hiroshi Sugimoto[1],
Jian-Ren Shen[7], Shinya Yoshikawa[3] & Hideo Ago[1]

We report a method of femtosecond crystallography for solving radiation damage–free crystal structures of large proteins at sub-angstrom spatial resolution, using a large single crystal and the femtosecond pulses of an X-ray free-electron laser (XFEL). We demonstrated the performance of the method by determining a 1.9-Å radiation damage–free structure of bovine cytochrome c oxidase, a large (420-kDa), highly radiation-sensitive membrane protein.

The modern synchrotron has contributed to high-resolution crystal structure determination by providing an especially bright X-ray beam. However, a method of radiation damage–free crystal structure analysis of proteins has not been fully realized. This is because reactive species, which form on a picosecond timescale upon X-ray irradiation, attack the molecules in the crystal even at cryogenic temperatures[1,2].

We report a radiation damage–free data collection method, developed at the SPring-8 angstrom compact free-electron laser (SACLA) facility and based on a combined usage of a large crystal and the femtosecond X-ray pulses of an XFEL. A femtosecond X-ray pulse provides a diffraction image before the chemical structure of the protein is destroyed by the X-ray pulse[3]. Early studies using photosystem II, a metalloprotein susceptible to such

radiation damage, showed that the femtosecond exposure time is markedly shorter than the time needed for the development of radiation damage[4,5].

In the serial femtosecond crystallography approach, a stream of microcrystals is flowed into the XFEL beam such that occasionally a microcrystal will be hit with an X-ray pulse and produce a diffraction pattern. The intensity of Bragg spots, each of which is regarded as a diffraction from a three-dimensional periodic structure, must be reconstructed in a Monte Carlo approach[6], which involves taking the average of many independent observations of the intensity of each Bragg spot from different microcrystals illuminated by XFEL pulses having variable pulse parameters. As the total diffraction count generally scales linearly with the crystal volume interacting with the X-ray, a large crystal is desirable for high-resolution data collection[7], even though there are potential drawbacks, such as extensive mosaicity and lattice defects. Large crystals also afford the capacity to record a series of diffraction images in a known rotation sequence, which allows accurate scaling and merging of data using existing tools developed for traditional crystallography instead of a Monte Carlo approach that requires highly redundant observation of the intensity of Bragg spots.

In our approach, we rotated a large single crystal with a size range of several hundred micrometers by a small angle in a stepwise fashion in order to record still diffraction images that are discrete and sequential in the course of the crystal rotation. The size of the stepwise crystal rotation, specifying how finely a diffraction spot is sampled along the crystal rotation, has a profound effect on the quality of the resultant diffraction data. We obtained preliminary experimental results using sequential still diffraction images of a lysozyme crystal collected at beamline BL26B2/SPring-8, finding that one-third of a crystal mosaicity was an optimal step size (Online Methods).

We recorded each diffraction image from an independent portion of the crystal, translating the crystal before every X-ray pulse irradiation so as to maintain a 50-µm distance apart from neighboring beam footprints (**Fig. 1a**). The translational size of 50 µm was experimentally determined to be a size free of radiation damage (Online Methods and **Fig. 1b**).

We applied the method to multiple crystals of bovine heart cytochrome c oxidase (CcO)[8], an O_2 reduction–coupled proton pump with a unit cell volume of 6.7 M$Å^3$, at experimental hutch 3 of beamline 3 of SACLA[9], to elucidate the radiation damage–free ligand structure of CcO in the fully oxidized 'resting fast' state

Figure 1 | Schematic of the data collection approach using femtosecond crystallography at SACLA. (**a**) Experimental setup for femtosecond crystallography characterized by the combined usage of a large crystal and femtosecond XFEL pulses. The distance between each irradiation point was set to 50 μm in the horizontal direction for any rotation angle. The vertical step size was varied depending on the rotation angle so as to maintain a 50-μm distance between two neighboring beam footprints. (**b**) Propagation of the radiation damage on the CcO crystal. Shown is the variation in the number of diffraction spots observed along the vertical direction, as defined in **Figure 1a**. The blue and red curves were measured with sufficiently attenuated XFEL pulses before and after a non-attenuated XFEL pulse irradiation (at 0 μm), respectively.
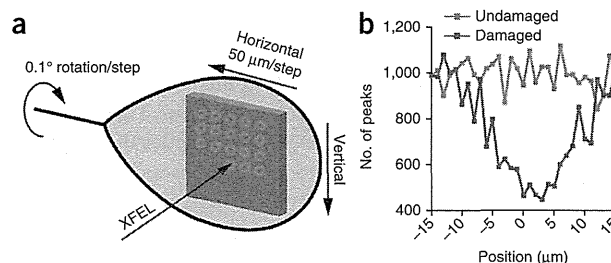
(hereafter called 'the oxidized state') (**Supplementary Note**). The O_2-reduction site of CcO[10] is composed of heme a_3 iron (Fe_{a3}) and copper (Cu_B) sites as well as a putative peroxide ligand in the oxidized state in the purified preparation used in this study[11]. X-ray irradiation reduces the peroxide ligand to give a hydroxide bound state (Fe_{a3}^{3+}-OH^-, Cu_B^{2+}-OH^-), resulting in a longer interatomic distance between the peroxide oxygen atoms than is found in a typical peroxide molecule[11].

We collected 1,396 still diffraction images, each rotated by a step of 0.1°, from 76 crystals of CcO at 100 K, each with a maximum width and thickness of 500 μm and 100 μm, respectively. The crystal rotation step of 0.1° was less than one-third of the average crystal mosaicity of 0.35°. The radiation dose was 9.9 MGy and was approximately the same for each exposure. The time between shots was 5 s according to the readout time of the detector. 1,107 images were processed and scaled together using MOSFLM[12] and SCALA[13] (**Supplementary Table 1**). The sequential still diffraction images, covering an angular range of approximately 1.5° on average (the maximum and minimum were approximately 10° and 0.5°, respectively) were collected from one crystal. The R_{merge} and R_{pim} (ref. 14) of the 1.9-Å diffraction data were 24.3% and 13.7%, respectively.

We evaluated radiation damage to the peroxide ligand as follows (Online Methods). An $F_o - F_c$ map calculated with the phases obtained by the refinement, performed without any atoms present between Fe_{a3} and Cu_B, possessed an elongated electron density shape (**Fig. 2a**). This is consistent with two oxygen atoms of a peroxide anion. The peak height of the water molecule bound to the Oγ of Tyr244 in the $F_o - F_c$ map (**Fig. 2a**), which increases along with the reduction of the peroxide ligand, is a reliable and sensitive way to determine the radiation damage to a peroxide

ligand[11]. The refined structure without the peroxide ligand obtained using XFEL radiation had a water peak (38% of the single water molecule) and intensity virtually identical to that of cyanide-bound oxidized CcO (37%). The water peak was most likely due to spontaneously introduced exogenous water molecules, as cyanide completely removes the bound peroxide. We compared our XFEL result to data collected on standard (BL44XU/SPring-8) and microfocus (BL32XU/SPring-8) synchrotron beamlines. The observed XFEL peak was clearly weaker than the intensities observed in the data obtained at BL44XU/SPring-8 (62%) and BL32XU/SPring-8 (43%) (**Supplementary Table 2**). The water-peak intensity clearly indicated the absence of any radiation damage to the bound peroxide ligand using our new XFEL-based method.

During structure model refinement without any restraint conditions on the peroxide oxygen atoms, the O-O distance converged to 1.75 Å. It was the shortest relative to those refined against the data collected at beamlines of BL44XU/SPring-8 in previous work[11] (1.90 Å) and BL32XU/SPring-8 (1.81 Å). We refined six structural models with different O-O distances constrained to 1.20 Å, 1.30 Å, 1.40 Å, 1.50 Å, 1.60 Å and 1.70 Å, and we evaluated their structures in terms of the residual electron density of their $F_o - F_c$ maps (**Supplementary Fig. 1**), as in the previous study[11]. Shortening the O-O distance increased the intensity of a residual peak close to Cu_B and near the peroxide, whereas a higher peak, located just next to the bond between the two oxygen atoms of the peroxide, appeared upon elongation of the O-O distance. The convergence point for the appearance of the two different residual densities with an increase in the O-O distance was between 1.50 Å and 1.60 Å for the XFEL data, whereas it was 1.70 Å and 1.65 Å for the data obtained at BL44XU/SPring-8 (ref. 11) and BL32XU/SPring-8, respectively. These results showed that the reduction of the peroxide was minimized by using XFEL femtosecond pulses.
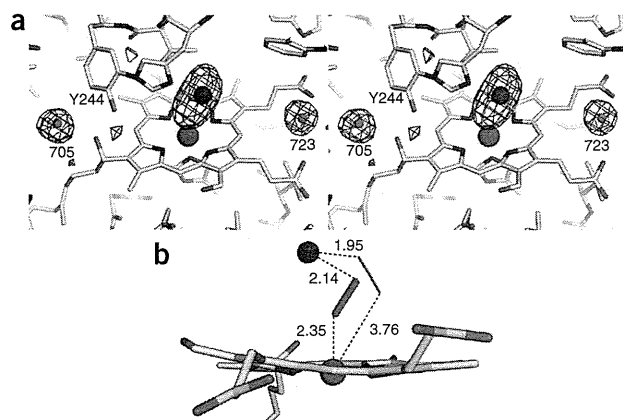
Figure 2 | The O_2-reduction site free of radiation damage. (**a**) $F_o - F_c$ map in stereo at 1.9-Å resolution for the XFEL data. Electron density cages at the 3σ level ($\sigma/0.041$ e$^-$/Å^3) are depicted along with the structural model of the O_2-reduction site, including the Fe_{a3} (red) and Cu_B (blue) atoms and the oxygen atoms (cyan) of the two water molecules. The two water molecules 705 and 723 were not included in the structural refinements in order to compare the electron densities with the electron density between Fe_{a3} and Cu_B as well as the electron density near Tyr244 Oγ. The peak height of both water molecules is 9.0σ, and the peak height near Tyr244 is 3.4σ. (**b**) Diagram of ligand binding. A stick model of the O_2-reduction site with a major component (95%; thick stick) and a minor component (5%; thin stick) of peroxide anions, both with an O-O distance of 1.55 Å. The colors used to designate the atom species are the same as in **a**. The numerical values indicated between the metal atoms and the oxygen atoms of the peroxide anions represent the distance in angstroms.

The appearance of residual electron densities beside but not on the O-O bond in the refinements with the constrained O-O distances of 1.60 or 1.70 Å (**Supplementary Fig. 1**) provides a strong indication of the existence of a minor component. The main component of the peroxide with 95% occupancy under the constraint of the O-O distance was refined to 1.55 Å and was located with a minor component of the peroxide with an occupancy of 5% on the residual densities (Online Methods). We judged from the residual electron density of the $F_o - F_c$ map at the binuclear center that the structural refinement converged well (**Supplementary Fig. 2a**). We performed similar calculations that included the minor components using other structural models with O-O distances of 1.60 Å and 1.50 Å, resulting in residual densities of the $F_o - F_c$ maps that were higher than that of the 1.55-Å model. Consequently, the present X-ray structural analysis using the XFEL data indicated that a peroxide anion with an O-O distance of 1.55 Å exhibited multiple conformations in the binuclear center, in which the main component, with 95% occupancy, had O-Cu_B and O-Fe_{a3} distances of 2.14 Å and 2.35 Å, respectively, and the minor component had analogous distances of 1.95 Å and 3.76 Å (**Fig. 2b**). The O-O distance obtained in this study was similar to that of a typical peroxide (Online Methods). The hydroxide bound state[11] (Fe_{a3}^{3+}-OH⁻, Cu_B^{2+}-OH⁻) caused by X-ray irradiation was undetectable, even as a minor component, in the X-ray structural data from the XFEL data. This indicated that our method afforded an undamaged structure. Further details on the ligand assignment are provided in the **Supplementary Results, Supplementary Figure 2b** and **Supplementary Note**.

Our method of using large single crystals for XFEL analysis yielded a radiation damage–free and high-resolution crystal structure of CcO, which could not be obtained using the microbeam of beamline BL32XU/SPring-8 at the same dose per image with a longer exposure time, or by a low-dose experiment at beamline BL44XU/SPring-8 (ref. 11). The ultrashort, femtosecond exposure time therefore provides an advantage over the use of a synchrotron beamline. The structure determination of redox-sensitive metalloproteins, such as photosystem II (ref. 15) and nitric oxide reductase[16], whose metal-ion reduction occurs below a dose for the diffraction intensity data collection using a synchrotron beamline, will especially benefit from this method. Because CcO is one of the proteins most sensitive to X-ray irradiation[11], a 50-μm separation between two irradiation points would be an appropriate displacement for most protein crystals.

METHODS
Methods and any associated references are available in the online version of the paper.

Accession codes. Protein Data Bank: the coordinate and the structure factor have been deposited with accession code 3WG7.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS
All authors contributed to the diffraction experiments at SACLA and discussed and commented on the results and the manuscript. K.H. and K.S.-I. planned experiments. K.T., Y.I., M. Yabashi and T.I. contributed to beamline operation. G.U., T.H., H.M. and M. Yamamoto contributed to development of the experimental instruments. N.Y., S.T., K.K., M.H., K.M., T.K. and E.Y. conducted protein purification, crystallization and diffraction experiments. S.Y. and T.T. wrote the manuscript. J.-R.S., T.O., H.S. and H.A. organized the damage-free diffraction experiment at SACLA.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Southworth-Davies, R.J., Medina, M.A., Carmichael, I. & Garman, E.F. *Structure* **15**, 1531–1541 (2007).
2. Yano, J. *et al. Proc. Natl. Acad. Sci. USA* **102**, 12047–12052 (2005).
3. Neutze, R., Wouts, R., van der Spoel, D., Weckert, E. & Hajdu, J. *Nature* **406**, 752–757 (2000).
4. Kern, J. *et al. Science* **340**, 491–495 (2013).
5. Davis, K.M. *et al. J. Phys. Chem. Lett.* **3**, 1858–1864 (2012).
6. White, T.A. *et al. J. Appl. Crystallogr.* **45**, 335–341 (2012).
7. Holton, J.M. & Frankel, K.A. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 393–408 (2010).
8. Tsukihara, T. *et al. Science* **269**, 1069–1074 (1995).
9. Ishikawa, T. *et al. Nat. Photonics* **6**, 540–544 (2012).
10. Yoshikawa, S., Muramoto, K. & Shinzawa-Itoh, K. *Annu. Rev. Biophys.* **40**, 205–223 (2011).
11. Aoyama, H. *et al. Proc. Natl. Acad. Sci. USA* **106**, 2165–2169 (2009).
12. Powell, H.R., Johnson, O. & Leslie, A.G.W. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1195–1203 (2013).
13. Evans, P. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
14. Evans, P.R. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 282–292 (2011).
15. Umena, Y., Kawakami, K., Shen, J.R. & Kamiya, N. *Nature* **473**, 55–60 (2011).
16. Hino, T. *et al. Science* **330**, 1666–1670 (2010).

ONLINE METHODS

Pulse parameters of SACLA. The pulse parameters of SACLA were as follows: pulse duration, 10 fs or shorter; pulse flux, 3.5×10^{10} photons; photon energy, 10.0 keV (1.241 Å); energy bandwidth, 0.6% (FWHM) (57 eV/10 keV); beam size, 1.8 (H) × 1.2 (V) μm^2. The parameters for shot-to-shot variation were as follows: photon energy: 0.07% (s.d.) (7 eV/10 keV); intensity, 12% (s.d.) (36 μJ/290 μJ at the upper stream beam monitor 1).

Evaluation of the propagation of the radiation damage induced by XFEL irradiation of CcO crystals. Evaluation of the damage propagation as a large-scale disorder in a CcO crystal cooled at 100 K was conducted along the vertical and horizontal directions, as defined in **Figure 1**. This information is a condition that precedes the avoidance of radiation damage at a smaller scale of interest, such as the movement of a few atoms. The number of the diffraction spots in the images was counted at various distances (1 μm apart) from the XFEL pulse irradiation point (0 μm). Each image was taken by a sufficiently attenuated (100-fold) XFEL pulse of 10 keV. As a control experiment, images were taken at the same points before the non-attenuated XFEL irradiation. Substantial propagation of the radiation damage was not detectable beyond 11 μm from the non-attenuated XFEL irradiation at 0 μm, as shown in **Figure 1**. Essentially the same results were obtained for the damage propagation in the horizontal direction. The distance between the irradiation points was set to 50 μm in the horizontal direction for each rotation angle. The translational distance in the vertical direction was varied along with the crystal rotation in order to maintain a 50-μm distance between neighboring beam footprints.

The validity of processing still diffraction images. Crystals were rotated by a small angle in a stepwise fashion to record sequential still diffraction images from a crystal. The appropriate angular step of crystal rotation enabled a sampling of the diffraction intensity of a diffraction spot using just a few still diffraction images, resulting in a few partial diffraction intensities that were discrete and sequential in the crystal rotation for each diffraction spot. Also, the programs designed for data processing of the oscillation images were able to process the still diffraction images by regarding each still diffraction image as an oscillation image with an oscillation angle equivalent to the crystal rotation step, even though some parts of the diffraction intensity of a given diffraction spot would be missed in the still data collection. The shot-to-shot variation in photon energy of 0.07% at 10 keV contributes to the stable data processing. The energy bandwidth of 0.6% (FWHM) at the photon energy of 10 keV (1.241 Å) may help ameliorate the problem derived from the still data collection in an analogous manner to the role of the pink beam in the pink Laue concept[17], although the actual effects of the current XFEL pulse on data collection need further investigation in future studies. The capacity to have still diffraction images that are discrete and sequential in a rotating crystal is a critical characteristic that makes the data collection possible, and it is distinctly different from the serial femtosecond crystallography performed using diffraction images that are completely independent of each other.

The step size of the crystal rotation, specifying how finely a diffraction spot along the crystal rotation is sampled, has a profound effect on the quality of the resultant diffraction data. To estimate the appropriate step width, we performed a preliminary experiment. A sequential data set of 1,800 still images by 0.05° steps was collected from a standard lysozyme crystal at beamline BL26B2/SPring-8. Three types of data processing with a conventional data processing program (MOSFLM[12]) were performed using all of the images, skipping one of two, and skipping three of four to simulate the step scan methods with a step width of 0.05°, 0.10° and 0.20°, respectively. The result is summarized in **Supplementary Table 3**, with the normal oscillation data set provided for comparison. The crystal mosaicity of the sample estimated from the normal data set was 0.42°.

The structural refinements of lysozyme against the three diffraction data sets corresponding to the step width of 0.05°, 0.10° and 0.20° were performed, in addition to the refinement using the normal diffraction data set collected by the conventional oscillation method with a 0.05° oscillation angle. The structural analysis and the refinement were performed by PHENIX[18]. The PDB accession ID of the starting model was 1LZC (ref. 19). **Supplementary Table 3** provides a summary of the results. All refinements converged to reasonable R and R_{free} values, although the statistical value of the diffraction data worsened along with an increase in the step width. The r.m.s. deviations of a superimposition of the crystal structure of the normal oscillation data on the structures refined against the still data were lower than 0.033 Å in all of the pairs, and the density correlation coefficient calculated by PHENIX[18] was higher than 0.897 in all of the refinements. These results show that the diffraction data corresponding to the rotation step of 0.20° as one-half of the crystal mosaicity of 0.42° are still available for structural refinement of a comparable quality to that using the normal oscillation data. The crystallographic refinement of lysozyme showed that half of the crystal mosaicity was of an acceptable step size. Therefore, one-third of the crystal mosaicity was determined to be a reliable measure.

Preparation of crystals of the oxidized bovine heart CcO. CcO in the oxidized state was purified from bovine heart mitochondria and crystallized as described previously[8,20]. The oxidized crystals were stabilized at 4 °C in stabilization solution (40 mM MES-Tris buffer, pH 5.8, containing 0.2% (w/v) n-decyl-β-D-maltoside and 1% (w/v) poly(ethylene glycol) 4,000 (Sigma)). To prepare the CN-bound oxidized crystals, we soaked the oxidized crystals in the CN solution, which consists of the stabilization solution plus 10 mM potassium cyanide. The CN solution was replaced by freshly prepared solution once per day. The crystals were frozen in a cryo-nitrogen stream at 100 K in the presence of 45% (v/v) ethylene glycol as a cryoprotectant.

X-ray diffraction experiments of CN-bound oxidized crystals of the bovine heart CcO. All of the X-ray experiments were carried out at beamline BL44XU/SPring-8. The beamline was equipped with an image plate detector, DIP6040. The X-ray beam cross-section for X-ray diffraction experiments was 70–80 μm × 50 μm, and the wavelength was 0.9 Å. Each 0.5° oscillation frame was taken by a 10-s exposure at 100 K. Data processing and scaling were carried out using DENZO and SCALEPACK[21]. Statistics of the intensity data are provided in **Supplementary Table 4**.

Control experiments using the microfocus beam of beamline BL32XU/SPring-8. The control experiment was performed using the microfocus beam of beamline BL32XU/SPring-8 at 100 K. The crystals of CcO were prepared in the same way as described above. The beam features of the beamline BL32XU/SPring-8 were as follows: X-ray energy, 10.047 keV; beam size, 1.5 (V) × 1.5 (H) μm^2 (FWHM); photon number at the sample position, 3.6×10^{10} photons per image; exposure time, 1.3 seconds per image. By a preliminary experiment analogous to the result in **Figure 1b**, the size of the area in which the number of diffraction spots decreased by a 1.3-s exposure of the above beam was 6.0 μm (V) and 4.5 μm (H). On the basis of the area size, the irradiation points were placed on a crystal so as to maintain a 10-μm interbeam distance and thus escape the radiation damage that occurred in the previous exposures. The ratio of the interbeam distance to the size of the area with reduced diffraction spots was almost double, and this is comparable to the ratio of the data collection at SACLA. The 367 diffraction images with an oscillation angle of 0.3° per image were processed using XDS[22] and AIMLESS[23], resulting in the 1.9-Å intensity data, as shown in **Supplementary Table 5**.

X-ray structural analysis. The initial phase angles of the structural factors up to 4.0-Å resolution were obtained by the molecular replacement (MR)[24] method using the fully oxidized structure previously determined at 1.8-Å resolution (PDB ID code 2DYR)[25]. The phases were extended to 1.9-Å resolution by density modification[26] coupled with noncrystallographic symmetry averaging using the CCP4 (ref. 27) program DM. As the damaged and undamaged structures should be essentially equivalent at 4.0-Å resolution, the low-resolution MR should not bias the data excessively. An electron density map with minimum or no model bias was calculated using Fourier coefficients of $F_o \times \exp(i\alpha_{MR/DM})$, where $\alpha_{MR/DM}$ is the resultant phase. The structural refinement initiated using X-PLOR[28] was followed by REFMAC[29]. The bulk solvent correction and the anisotropic scaling of the observed and calculated structure amplitudes were incorporated into the refinement calculation. The anisotropic temperature factors for the iron, copper and zinc atoms were imposed on the calculated structure factors. The quality of the structural refinement was examined by the R and R_{free} values. There were no substantial structural changes in comparison to the crystal structure used as the search model in the MR. The two water molecules 705 and 723 shown in **Figure 2a** in the vicinity of the O_2-reduction site were not included in the structural refinement in order to compare their peak heights as a reference for the unknown groups in the $F_o - F_c$ maps. All of the $F_o - F_c$ maps were calculated with the Fourier coefficients and phases obtained by REFMAC[29] in the structural refinement.

Assessment of radiation damage of the diffraction data measured at SACLA. The peak height of the water molecule hydrogen-bonded to the Oγ of Tyr244 can be used as one criterion for the radiation damage of the bound peroxide. It has been shown by measurements performed at SPring-8 that the X-ray irradiation of the oxidized CcO crystal cooled at 100 K decreases the electron density of the peroxide ligand concomitantly with an increase in the peak height of the water molecule hydrogen-bonded to Oγ of

Tyr244 in the $F_o - F_c$ map[11]. Tyr244 functions as a scavenger of the water molecules produced by the X-ray irradiation–derived reduction of the peroxide bound at the O_2-reduction site. Thus, the peak intensity of the water molecule site, normalized by comparison with the average of the peak heights of the two water molecules near the O_2-reduction site (which are insensitive to X-ray irradiation), can be used as one criterion of radiation damage of a peroxide ligand.

To assess the degree of radiation damage, we performed the above described X-ray structural analysis and refinement without the models of the peroxide ligand and the two reference waters on each of four different diffraction data set: one was the diffraction data set measured at SACLA (this work), another was the diffraction data set collected at beamline BL44XU/SPring-8 with a dose of 0.33 MGy per image[11], a third was the diffraction data set collected at 100 K at beamline BL32XU/SPring-8 with the same dose as the data of SACLA (**Supplementary Table 5**), and the last was the diffraction data set on the crystal of cyanide-bound CcO measured at 100 K at BL44XU/SPring-8 (**Supplementary Table 4**). Cyanide completely replaces the peroxide ligand at the O_2-reduction site[30], and the water that becomes trapped by Tyr244 is not formed by the X-ray–induced reduction of the ligand; therefore, the peak height of the electron density can be used as the standard peak height free of radiation damage of the peroxide ligand. The peak heights of the water molecules are summarized for three different data sets in **Supplementary Table 2**.

Refinement of the O_2-reduction site. Three different structural refinements with two oxygen atoms between the two metal ions were performed to determine the precise structure of the O_2-reduction site. The procedure from the low-resolution MR to the structural refinement was described above. The first refinement was performed without any restraint for the O-O structure. Second, the refinements with constrained O-O bond distances of 1.20 Å, 1.30 Å, 1.40 Å, 1.50 Å, 1.60 Å and 1.70 Å were performed using the data measured at SACLA (this work) and beamlines of BL44XU/SPring-8 (ref. 11) and BL32XU/SPring-8. In the constrained refinement, the validity of the bond distances was evaluated using $F_o - F_c$ maps (**Supplementary Fig. 1**). Third, structural refinements using multiple conformations with different O-O distances were performed to reduce the residual densities between the two metal ions observed in the second refinement. For peroxide models with O-O distances of 1.50 Å, 1.55 Å and 1.60 Å, two oxygen atoms of a minor component having the same O-O distance as that of the main component were identified on the residual density sites. One of these was close to Cu_B, and the other was adjacent to the peroxide anion. All the parameters of the minor component and the bond distance of the major component were fixed during the refinement process. The statistical data on the structure determination for the peroxide model with an O-O distance of 1.55 Å are given in **Supplementary Table 1**.

According to the Cambridge Structural Database, the average O-O distance among 97 non-protein-derived peroxide structures that bridge two metal ions is 1.44 ± 0.06 Å (± s.d.). The distance range is between 1.29 Å and 1.59 Å, and 11 of the 97 distances are between 1.50 Å and 1.59 Å.

17. Dejoie, C. *et al. J. Appl. Crystallogr.* **46**, 791–794 (2013).
18. Adams, P.D. *et al. Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
19. Maenaka, K. *et al. J. Mol. Biol.* **247**, 281–293 (1995).
20. Mochizuki, M. *et al. J. Biol. Chem.* **274**, 33403–33411 (1999).
21. Otwinowski, Z. & Minor, W. *Methods Enzymol.* **276**, 307–326 (1997).
22. Kabsch, W. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
23. Evans, P.R. & Murshudov, G.N. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1204–1214 (2013).
24. Rossmann, M.G. & Blow, D.M. *Acta Crystallogr.* **15**, 24–31 (1962).
25. Shinzawa-Itoh, K. *et al. EMBO J.* **26**, 1713–1725 (2007).
26. Wang, B.C. *Methods Enzymol.* **115**, 90–112 (1985).
27. Collaborative Computational Project, Number 4. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
28. Brünger, A.T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
29. Murshudov, G.N. *et al. Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
30. Sakaguchi, M., Shinzawa-Itoh, K., Yoshikawa, S. & Ogura, T. *J. Bioenerg. Biomembr.* **42**, 241–243 (2010).

LETTER

Structural basis of Sec-independent membrane protein insertion by YidC

Kaoru Kumazaki[1,2]*, Shinobu Chiba[3]*, Mizuki Takemoto[1,2], Arata Furukawa[4], Ken-ichi Nishiyama[5], Yasunori Sugano[4], Takaharu Mori[6], Naoshi Dohmae[2], Kunio Hirata[7], Yoshiko Nakada-Nakura[8], Andrés D. Maturana[9], Yoshiki Tanaka[4], Hiroyuki Mori[10], Yuji Sugita[6], Fumio Arisaka[11], Koreaki Ito[3], Ryuichiro Ishitani[1,2], Tomoya Tsukazaki[4,12] & Osamu Nureki[1,2]

Newly synthesized membrane proteins must be accurately inserted into the membrane, folded and assembled for proper functioning. The protein YidC inserts its substrates into the membrane, thereby facilitating membrane protein assembly in bacteria; the homologous proteins Oxa1 and Alb3 have the same function in mitochondria and chloroplasts, respectively[1,2]. In the bacterial cytoplasmic membrane, YidC functions as an independent insertase and a membrane chaperone in cooperation with the translocon SecYEG[3–5]. Here we present the crystal structure of YidC from *Bacillus halodurans*, at 2.4 Å resolution. The structure reveals a novel fold, in which five conserved transmembrane helices form a positively charged hydrophilic groove that is open towards both the lipid bilayer and the cytoplasm but closed on the extracellular side. Structure-based *in vivo* analyses reveal that a conserved arginine residue in the groove is important for the insertion of membrane proteins by YidC. We propose an insertion mechanism for single-spanning membrane proteins, in which the hydrophilic environment generated by the groove recruits the extracellular regions of substrates into the low-dielectric environment of the membrane.

The Sec translocon, a protein-conducting channel conserved in all three phylogenetic domains, translocates secretory proteins across the membrane and inserts membrane proteins into the membrane by an hourglass-shaped pore formed by ten transmembrane helices[6–8]. In bacteria, another membrane protein, YidC, is involved in the folding and insertion of many membrane proteins[9–15], such as subunit c of the F_0F_1-ATPase (F_0c), and thus is essential for cell viability[3–5]. YidC is considered to function as both a Sec-dependent membrane chaperone and a Sec-independent insertase. In the Sec-independent pathway, YidC directly interacts with the translating ribosome to mediate the insertion of several single-spanning or double-spanning membrane proteins[16,17]. Previous studies have suggested that YidC forms a face-to-face dimer, creating a channel at its dimer interface[16,18]. By contrast, another research group has reported that a monomer of membrane-embedded YidC is sufficient for binding to the substrate-translating ribosome[19]. Oxa1, the mitochondrial homologue of YidC, reportedly functions as a voltage-gated membrane channel, as well as a membrane protein insertase, probably by forming a tetramer[20]. However, the lack of a high-resolution structure of YidC has limited our understanding of the molecular mechanism of YidC-mediated membrane protein insertion.

Members of the genus *Bacillus* have two *yidC* genes, encoding YidC1 and YidC2. We determined the crystal structures of two constructs of *B. halodurans* YidC2 ($YidC_{27–266}$ and $YidC_{27–267}$), which lack both the amino-terminal signal and the carboxy-terminal non-conserved sequences

(Fig. 1, Extended Data Table 1 and Extended Data Figs 1 and 2). A similar deletion variant of *Bacillus subtilis* YidC1 (also known as SpoIIIJ), an orthologue of *B. halodurans* YidC2 with 49.6% sequence identity, has comparable activity to the full-length protein *in vivo* (K248stop in Fig. 2a). Although the crystallographic asymmetric units of $YidC_{27–266}$ and $YidC_{27–267}$ contain one and two YidC molecules, respectively, the molecules in the crystalline lattice do not appear to form an effective oligomer, such as a face-to-face dimer[16,18] (Extended Data Fig. 3a, b). Recent fluorescence correlation spectroscopy and cryo-electron microscopy analyses of YidC bound to a translating ribosome showed that YidC exists as a monomer in detergent solution and in lipid membranes[17,19]. Consistent with these findings, our analysis using size exclusion chromatography coupled to multi-angle laser light scattering (SEC-MALLS) also showed that YidC exists as a monomer in detergent solution (Extended Data Fig. 3c–e). Because the overall structures of $YidC_{27–266}$ and $YidC_{27–267}$ are nearly identical (with a root mean square deviation of 1.79 Å over residues 27–266), we mainly describe the structure of $YidC_{27–266}$.

The $YidC_{27–266}$ structure consists of the N-terminal E1 region, the C-terminal C3 tail and the core region: the core region is composed of transmembrane helices 1–5 (TM1–5), connected by two cytoplasmic (C1 and C2) regions and two extracellular (E2 and E3) regions (Fig. 1c). The E1 region consists of the EH1 helix, whereas the C1 region forms a hairpin-like structure composed of two helices (CH1 and CH2) connected by a short loop. The residues in the C2 region and the C3 tail are structurally disordered. The EH1, CH1 and CH2 helices protrude from the core region and lie nearly parallel to the plane of the membrane. The EH1 helix is amphipathic, whereas the CH1 and CH2 helices contain mainly hydrophilic residues. This observation suggested that one face of the EH1 helix is embedded in the membrane, whereas the CH1 and CH2 helices are exposed to the solvent (Fig. 1a). This idea is also consistent with theoretical calculations using an implicit membrane model (see Methods).

A comparison of the structures of $YidC_{27–266}$ and $YidC_{27–267}$ suggested that the C1 region, which does not interact with the remainder of YidC, is flexible (Extended Data Fig. 4a–c). The CH1 helix forms a continuous helix with the TM1 helix, which is kinked at the conserved proline residues P78 and P94 (Fig. 1c). The proline residues and the partially disordered flexible loop connecting the CH2 and TM2 helices (residues 130–140), which is likely to be embedded in the membrane, could enhance the flexibility of the C1 region. Indeed, the C1 region has higher B-factors than the other regions (Extended Data Fig. 5a). A molecular dynamics simulation also suggested that the position of the C1 region fluctuates greatly in the lipid bilayer environment (Extended Data Fig. 6a, b).

[1]Department of Biological Sciences, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. [2]Global Research Cluster, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. [3]Faculty of Life Sciences, Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku, Kyoto 603-8555, Japan. [4]Department of Systems Biology, Graduate School of Biological Sciences, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan. [5]Cryobiofrontier Research Center, Faculty of Agriculture, Iwate University, 3-18-8 Ueda, Morioka, Iwate 020-8550, Japan. [6]Theoretical Molecular Science Laboratory, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan. [7]SR Life Science Instrumentation Unit, RIKEN SPring-8 Center, 1-1-1 Kouto, Sayo-cho, Sayo-gun, Hyogo 679-5148, Japan. [8]Department of Cell Biology, Graduate School of Medicine, Kyoto University, Yoshidakonoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. [9]Department of Bioengineering Sciences, Graduate School of Bioagricultural Sciences, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan. [10]Institute for Virus Research, Kyoto University, Shogoin Kawara-cho, Sakyo-ku, Kyoto 606-8507, Japan. [11]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa 226-8503, Japan. [12]JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.
*These authors contributed equally to this work.
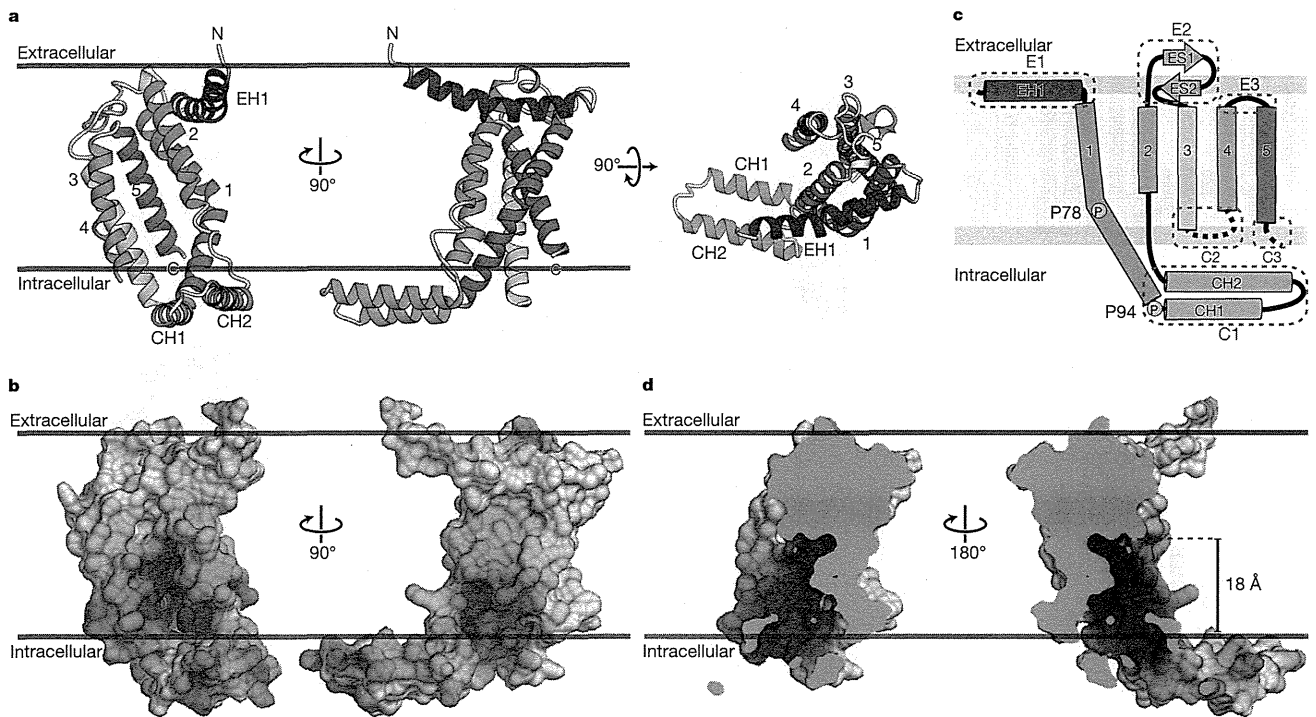
Figure 1 | Overall structure of YidC. a, Cartoon representations of *Bacillus halodurans* YidC viewed transversely through the membrane (left) and from the extracellular fluid (right). The structure is coloured from blue to red from the N to the C terminus. **b**, Surface model representations of YidC coloured according to electrostatic potential, ranging from blue ($+20\,kT/e$) to red

($-20\,kT/e$), where kT is thermal energy and e is the elementary charge. **c**, Topology diagram of YidC, coloured as in **a**. The grey area represents the cytoplasmic membrane. **d**, Cut-away molecular surface representations, viewed transversely through the membrane, coloured as in **b**.

To further investigate the functional importance of the C1 region, we performed a structure-based genetic analysis of *B. subtilis* SpoIIIJ. The membrane insertion of MifM, a single-spanning membrane protein, is mediated by the SpoIIIJ pathway[21]. In this analysis, we measured the β-galactosidase (LacZ) activity of a YidC2–LacZ translational fusion protein: this activity increased in response to a defect in the SpoIIIJ-dependent insertion of MifM (Extended Data Figs 7a and 8). Two C1-deletion mutants (Extended Data Fig. 7b) exhibited higher LacZ activities than wild-type SpoIIIJ (Fig. 2a), suggesting that the C1 region is crucial for the YidC-mediated membrane insertion of MifM. We then performed a growth complementation analysis using *B. subtilis* (as described in Extended Data Fig. 7c), which also suggested that the C1 region is important for SpoIIIJ activity (Fig. 2b).

The extracellular halves of the TM1–5 helices, in addition to the E2 region, are tightly packed with hydrophobic side chains (Extended Data Fig. 5b, c), whereas the cytoplasmic halves of the TM1–5 helices loosely interact with each other to form a groove ($\sim$2,000 Å^3). This groove contains many hydrophilic residues, including the conserved T68, R72, Q82, Q142, Q187, N248 and Q254 residues, and thus generates a hydrophilic environment in the lipid bilayer (Fig. 2c and Extended Data Fig. 1). The conserved R72 residue, the only charged residue in this groove, protrudes into the centre of the groove and creates a strong positive electrostatic potential in the groove (Fig. 1b, d). This hydrophilic groove is open to both the cytoplasmic side and the membrane interior. By contrast, the groove is sealed towards the extracellular side by the hydrophobic core and is not accessible from that side (Fig. 1d). The hydrophobic core, which consists of the hydrophobic residues in the extracellular half of YidC, has lower B-factors than the other regions, suggesting that it is rigid (Extended Data Fig. 5). Structure-based genetic analyses suggested the importance of this hydrophobic core for the function of YidC (Fig. 2a–c and Extended Data Fig. 7b). A comparison of the YidC$_{27-266}$ and YidC$_{27-267}$ structures revealed that the groove in YidC$_{27-267}$ is narrower

than that in YidC$_{27-266}$, suggesting that the groove has structural flexibility (Extended Data Fig. 4d). Molecular dynamics simulations also suggested that the size of the hydrophilic groove slightly fluctuated during a 1,000-ns simulation (Extended Data Fig. 6a). By contrast, the overall architecture of the core region, and the structure of the hydrophilic groove, remained stable in the lipid bilayer (Extended Data Fig. 6b), and the groove was constantly filled with $\sim$20 water molecules (Extended Data Fig. 6c). The extracellular side of the groove remained sealed by the hydrophobic protein residues and the aliphatic lipid chains and thus was impermeable to ions and water molecules during molecular dynamics simulations. Taking these findings together, YidC appears to provide a flexible hydrophilic groove in the membrane that is open towards both the cytoplasmic side and the membrane interior but tightly sealed on the extracellular side.

To investigate the functional importance of the hydrophilic groove, we examined the membrane insertion activities of SpoIIIJ molecules in which R73 (R72 in YidC; Fig. 2c) or the conserved glutamine residues had been mutated, by using *in vivo* genetic analyses. All tested mutations of R73 abolished MifM insertion activity, except for replacement of arginine with lysine (R73K), which slightly decreased the insertion activity (Fig. 2d). None of the R73 mutants, except for R73K, complemented the growth of wild-type SpoIIIJ- and wild-type YidC2-depleted cells (Fig. 2b). By contrast, mutation of Q83, Q140, Q187 or Q238 (Q82, Q142, Q187 and Q254, respectively, in YidC; Fig. 2c) to alanine did not affect MifM insertion activity (Fig. 2e). A similar result was obtained for a chimaera of the Pf3 coat protein, another single-spanning membrane protein inserted by the Sec-independent pathway, and the cytoplasmic region of MifM (Fig. 2f and Extended Data Fig. 9a). Taken together, these findings highlight the importance of the positive charge in the groove for insertion of MifM and the Pf3 coat protein into the membrane by YidC, whereas the conserved polar residues are probably important for creating the hydrophilic environment in the groove.
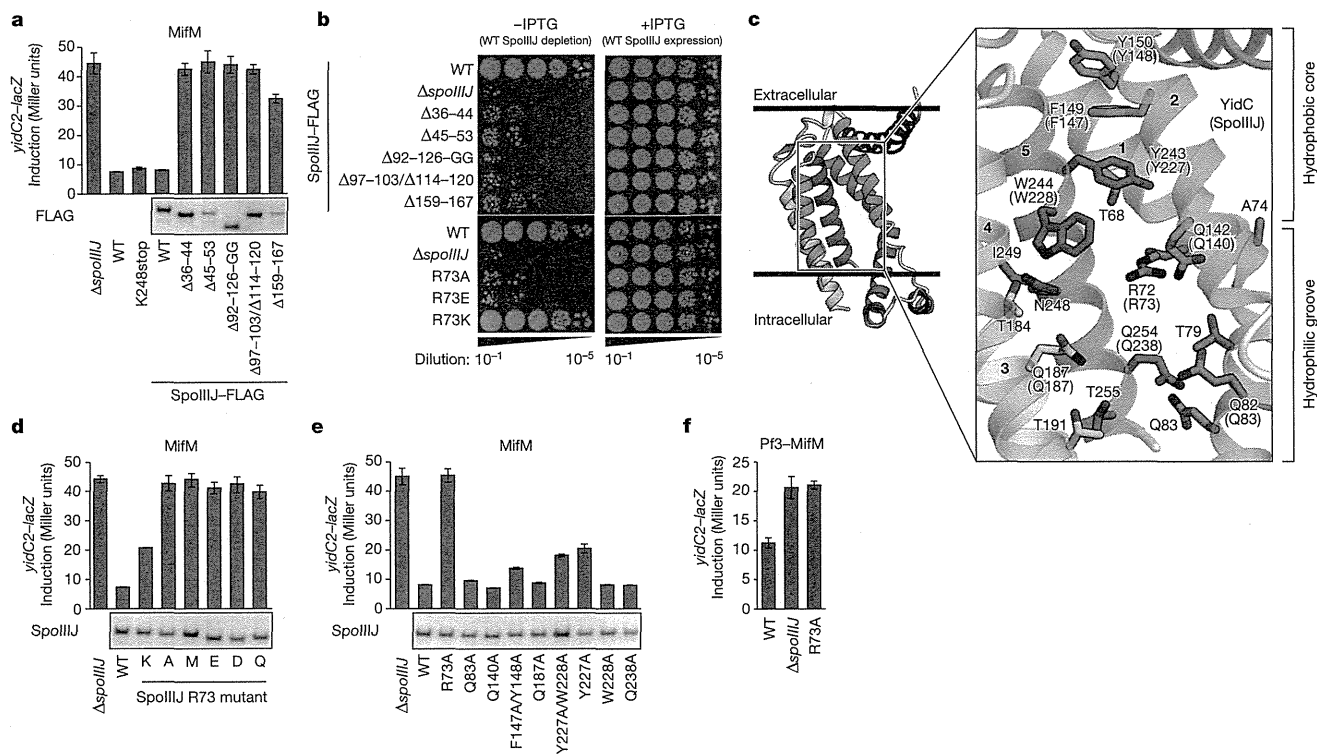
Figure 2 | In vivo functional analyses of YidC. a, d–f, Efficiencies of MifM and Pf3–MifM fusion protein insertion into the membrane, as determined by LacZ activity (mean ± s.d., n = 3) in *spoIIIJ* mutant cells, as indicated (top). The accumulation of these proteins in the cell is shown, detected by anti-FLAG or anti-SpoIIIJ immunoblotting (bottom). K248stop represents a *spoIIIJ* derivative that has a stop codon at position 248. Δ92–126-GG represents a mutant in which the entire C1 region has been replaced by a glycine–glycine

linker. Δ97–103/Δ114–120 represents a mutant in which both the CH1 and CH2 helices have been shortened by seven residues. **b,** Growth complementation of *Bacillus subtilis* cells reliant on a chromosomal SpoIIIJ mutant in the absence of isopropyl-β-D-thiogalactoside (IPTG). **c,** Close-up view of the hydrophilic groove of *Bacillus halodurans* YidC showing the side chains of the indicated residues. The corresponding residues in *B. subtilis* SpoIIIJ are indicated in parentheses. WT, wild type.
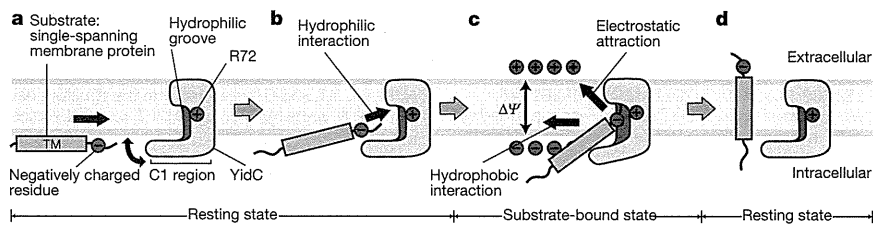
Several single-spanning membrane proteins, including MifM and the Pf3 coat protein, have acidic residues in their N-terminal extracellular tails. Thus, our results suggested that these acidic residues might interact with the arginine residue in the hydrophilic groove of YidC. To address this possibility, we examined the importance of these acidic residues by using an *in vivo* genetic analysis (Extended Data Fig. 9a). Mutating the acidic residues in the Pf3 coat protein had a less pronounced effect than mutating these residues in MifM; however, these mutations negatively affected the membrane insertion efficiencies of both MifM and the Pf3 coat protein, thereby supporting our hypothesis (Extended Data Fig. 9b). To further confirm that direct interactions occur between the substrate and the hydrophilic groove, we performed an *in vivo* site-directed ultraviolet-radiation crosslinking analysis, using p-benzoyl-L-phenylalanine (pBpa)[22]. pBpa was introduced at Q187 or W244 in the groove and at A74 or I249 on the exterior surface of YidC (Fig. 2c and Extended Data Fig. 1). *Escherichia coli* cells coexpressing MifM and the pBpa variant of YidC were exposed to ultraviolet radiation, and products that were crosslinked with MifM were analysed by immunoblotting. We detected crosslinked products only at positions 187 and 244, suggesting that the groove can interact with MifM (Fig. 3). Taken together, these observations suggested that the site around R72 in the hydrophilic groove is the substrate-binding site, through recognition of the hydrophilic residues of the substrate such as the acidic residues in MifM and the Pf3 coat protein.

On the basis of these structural and functional analyses, we propose a mechanism for the insertion of single-spanning membrane proteins with an acidic N-terminal extracellular region, such as MifM and the Pf3 coat protein, that is mediated by monomeric YidC (Fig. 4). In this mechanism, the substrate protein initially interacts with the C1 region (Fig. 4a) and is then transiently captured in the hydrophilic groove of YidC. This substrate binding can induce structural changes in the hydrophilic groove,

to accommodate the various substrate proteins. In this context, the conserved arginine residue in the groove might participate in substrate recognition (Fig. 4b). Subsequently, the transmembrane region of the substrate
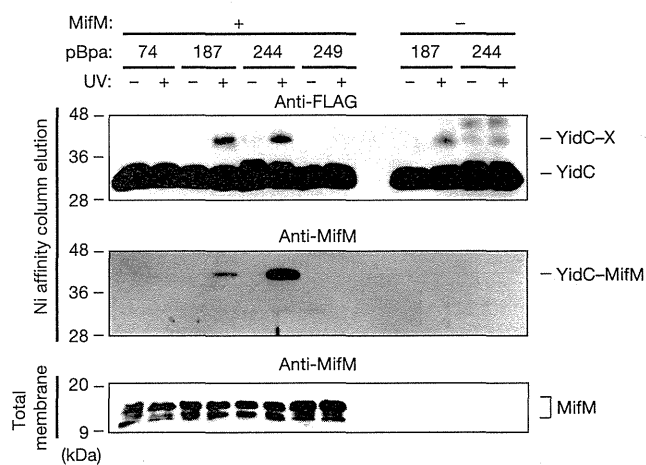
Figure 3 | Substrate binding to the hydrophilic groove. *In vivo* photo-crosslinking between the hydrophilic groove of YidC and MifM. The amino acid positions at which pBpa was introduced into YidC are indicated. Membrane proteins were purified by separation on a nickel affinity column and analysed by SDS–PAGE, and YidC–FLAG–His$_8$ and MifM were detected by anti-FLAG and anti-MifM immunoblotting, respectively. The ~40 kDa-bands are attributed to products in which YidC is crosslinked to MifM and to unidentified endogenous proteins, as indicated by YidC–X. The accumulation of MifM in the membrane is shown (bottom).

Figure 4 | Proposed model for membrane insertion of a single-spanning membrane protein. The obtained crystal structure probably represents the resting state before substrate binding. The hydrophilic region of the substrate would then be transiently captured in the hydrophilic groove of YidC (**a, b**), resulting in the substrate-bound state (**c**). Substrate release into the membrane can be facilitated by the hydrophobic interaction between the transmembrane (TM) region and the lipid aliphatic chains, as well as by the membrane potential ($\Delta\Psi$), attracting the negatively charged residues of the extracellular region of the single-spanning membrane protein by electrostatic attraction (**c, d**).

protein is released into the membrane, with the hydrophilic residues translocated to the extracellular side. The substrate release can be facilitated by the hydrophobic interaction between the transmembrane region and the lipid aliphatic chains. The membrane potential could also facilitate this process, by attracting the negative charge of the extracellular tail[23,24] (Fig. 4c).

The above mechanism clearly explains the insertion of a certain class of membrane proteins: that is, single-spanning membrane proteins with an acidic N-terminal extracellular region. However, it cannot account for the YidC-mediated insertion of other classes of membrane protein. For example, for *E. coli* YidC, deletion of the C1 region or simultaneous substitution of five amino acids, including the conserved arginine residue, with serine, was not found to impair the insertion of an M13 pro-coat derivative, a double-spanning membrane protein[25]. Therefore, it is likely that these membrane proteins are inserted by a different mechanism, which might involve either sites in the hydrophilic groove other than the arginine residue or transient oligomer formation by YidC. Further structural and biological studies are required to clarify the mechanism of the YidC-mediated insertion of these membrane protein classes.

METHODS SUMMARY

Histidine-tagged *Bacillus halodurans* YidC2 (YidC) was overproduced in *Escherichia coli*, solubilized from the membrane with *n*-dodecyl-β-D-maltoside and cholesteryl hemisuccinate and purified by successive nickel (Ni-NTA) chromatography and gel filtration chromatography steps. The histidine tag was then cleaved with tobacco etch virus (TEV) protease. Crystals were grown in a lipidic cubic phase, using monoolein. Diffraction data were collected on beamline BL32XU at SPring-8. The structure of YidC was determined by the multi-wavelength anomalous diffraction method, using a methyl-mercury-chloride-derivatized YidC (Y150C mutant) crystal and was refined to an R_{work}/R_{free} of 24.2%/25.9% at 2.4 Å resolution. Molecular dynamics simulations of YidC in the explicit phosphoryl oleoyl phosphatidylethanolamine (POPE) lipid bilayer were performed with the NAMD program (version 2.8) for 1,000 ns. The MifM insertion activity of YidC was analysed by using a MifM-based assay in which the β-galactosidase (LacZ) activity of a YidC2–LacZ translational fusion increases in response to a defect in the SpoIIIJ-dependent insertion of MifM. The growth complementation assay was performed using mutant *Bacillus subtilis* cells, in which the *yidC2* gene on the chromosome was disrupted and the *spoIIIJ* gene was mutated, with rescue by a plasmid harbouring the wild-type isopropyl-β-D-thiogalactoside (IPTG)-inducible *spoIIIJ* gene. *In vivo* photo-crosslinking was performed using pBpa. *B. subtilis* MifM and pBpa variants of YidC were overexpressed in *E. coli*, and the cells were exposed to ultraviolet radiation. YidC was then purified by Ni-NTA chromatography, and products that were crosslinked with MifM were detected by immunoblotting. See Methods for further detail.

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper.

1. Funes, S., Kauff, F., van der Sluis, E. O., Ott, M. & Herrmann, J. M. Evolution of YidC/Oxa1/Alb3 insertases: three independent gene duplications followed by functional specialization in bacteria, mitochondria and chloroplasts. *Biol. Chem.* **392**, 13–19 (2011).

2. Saller, M. J., Wu, Z. C., de Keyzer, J. & Driessen, A. J. M. The YidC/Oxa1/Alb3 protein family: common principles and distinct features. *Biol. Chem.* **393**, 1279–1290 (2012).
3. Samuelson, J. C. *et al.* YidC mediates membrane protein insertion in bacteria. *Nature* **406**, 637–641 (2000).
4. Scotti, P. A. *et al.* YidC, the *Escherichia coli* homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *EMBO J.* **19**, 542–549 (2000).
5. Dalbey, R. E., Wang, P. & Kuhn, A. Assembly of bacterial inner membrane proteins. *Annu. Rev. Biochem.* **80**, 161–187 (2011).
6. Park, E. & Rapoport, T. A. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu. Rev. Biophys.* **41**, 21–40 (2012).
7. van den Berg, B. *et al.* X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004).
8. Tsukazaki, T. *et al.* Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature* **455**, 988–991 (2008).
9. Nagamori, S., Smirnova, I. N. & Kaback, H. R. Role of YidC in folding of polytopic membrane proteins. *J. Cell Biol.* **165**, 53–62 (2004).
10. Wagner, S. *et al.* Biogenesis of MalF and the MalFGK$_2$ maltose transport complex in *Escherichia coli* requires YidC. *J. Biol. Chem.* **283**, 17881–17890 (2008).
11. Yi, L. *et al.* YidC is strictly required for membrane insertion of subunits a and c of the F$_1$F$_0$ATP synthase and SecE of the SecYEG translocase. *Biochemistry* **42**, 10537–10544 (2003).
12. du Plessis, D. J. F., Nouwen, N. & Driessen, A. J. M. Subunit a of cytochrome o oxidase requires both YidC and SecYEG for membrane insertion. *J. Biol. Chem.* **281**, 12248–12252 (2006).
13. Price, C. E. & Driessen, A. J. M. Conserved negative charges in the transmembrane segments of subunit K of the NADH:ubiquinone oxidoreductase determine its dependence on YidC for membrane insertion. *J. Biol. Chem.* **285**, 3575–3581 (2010).
14. Serek, J. *et al. Escherichia coli* YidC is a membrane insertase for Sec-independent proteins. *EMBO J.* **23**, 294–301 (2004).
15. Facey, S. J., Neugebauer, S. A., Krauss, S. & Kuhn, A. The mechanosensitive channel protein MscL is targeted by the SRP to the novel YidC membrane insertion pathway of *Escherichia coli. J. Mol. Biol.* **365**, 995–1004 (2007).
16. Kohler, R. *et al.* YidC and Oxa1 form dimeric insertion pores on the translating ribosome. *Mol. Cell* **34**, 344–353 (2009).
17. Seitl, I., Wickles, S., Beckmann, R., Kuhn, A. & Kiefer, D. The C-terminal regions of YidC from *Rhodopirellula baltica* and *Oceanicaulis alexandrii* bind to ribosomes and partially substitute for SRP receptor function in *Escherichia coli. Mol. Microbiol.* **91**, 408–421 (2014).
18. Lotz, M., Haase, W., Kühlbrandt, W. & Collinson, I. Projection structure of yidC: a conserved mediator of membrane protein assembly. *J. Mol. Biol.* **375**, 901–907 (2008).
19. Kedrov, A. *et al.* Elucidating the native architecture of the YidC:ribosome complex. *J. Mol. Biol.* **425**, 4112–4124 (2013).
20. Krüger, V. *et al.* The mitochondrial oxidase assembly protein1 (Oxa1) insertase forms a membrane pore in lipid bilayers. *J. Biol. Chem.* **287**, 33314–33326 (2012).
21. Chiba, S., Lamsa, A. & Pogliano, K. A ribosome-nascent chain sensor of membrane protein biogenesis in *Bacillus subtilis. EMBO J.* **28**, 3461–3475 (2009).
22. Mori, H. & Ito, K. Different modes of SecY–SecA interactions revealed by site-directed *in vivo* photo-cross-linking. *Proc. Natl Acad. Sci. USA* **103**, 16159–16164 (2006).
23. Chen, M. *et al.* Direct interaction of YidC with the Sec-independent Pf3 coat protein during its membrane protein insertion. *J. Biol. Chem.* **277**, 7670–7675 (2002).
24. Zhu, L. L., Wasey, A., White, S. H. & Dalbey, R. E. Charge-composition features of model single-span membrane proteins that determine selection of YidC and SecYEG translocase pathways in *Escherichia coli. J. Biol. Chem.* **288**, 7704–7716 (2013).
25. Jiang, F. *et al.* Defining the regions of *Escherichia coli* YidC that contribute to activity. *J. Biol. Chem.* **278**, 48965–48972 (2003).

Author Contributions K.K. performed the crystallization and structure determination. S.C. performed the genetic analyses. K.K., A.F., K.-I.N., Y. Sugano, A.D.M., Y.T., H.M. and T.T. performed the functional analysis. M.T., T.M., Y. Sugita and R.I. performed the molecular dynamics simulation. K.K., N.D. and F.A. identified the molecular mass. K.H., Y.N.-N., R.I., T.T. and O.N. assisted with the structure determination. K.K., S.C., K.I., R.I., T.T. and O.N. wrote the manuscript. T.T. and O.N. directed and supervised all of the research.

Author Information The atomic coordinates and structure factors for $YidC_{27-266}$ and $YidC_{27-267}$ have been deposited in the Protein Data Bank under accession numbers 3WO6 and 3WO7, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to O.N. (nureki@bs.s.u-tokyo.ac.jp) or T.T. (ttsukaza@bs.naist.jp).

METHODS

Cloning, expression and purification of YidC for structure determination. *Bacillus halodurans* YidC2 (YidC) was cloned into a pET-modified vector[26]. The two resultant plasmids, encoding YidC$_{1-26}$-H$_8$-ENLYFQGQ-YidC$_{27-266}$ (YidC$_{27-266}$) and YidC$_{1-26}$-ENLYFQGQ-YidC$_{27-267}$-LESSV-ENLYFQGQ-green fluorescent protein (GFP)-H$_8$ (YidC$_{27-267}$), were expressed in *Escherichia coli* C41(DE3) cells harbouring pRARE (Novagen), and the proteins were purified using the following protocol. The cells were grown in a 5 l culture at 37 °C to an absorbance at 600 nm of 0.7, and expression was induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG) at 15 °C for 16 h. The cells were then harvested by centrifugation (4,500 g, 10 min, 4 °C), resuspended in buffer containing 20 mM Tris-HCl, pH 8.0, and 0.1 mM phenylmethyl-sulphonyl fluoride, and disrupted by two passages through a Microfluidizer processor (Microfluidics) at 15,000 p.s.i. After removal of the debris by centrifugation (12,000 g, 30 min, 4 °C), the supernatant was ultracentrifuged (138,000 g, 1 h, 4 °C) to pellet the membranes, which were then solubilized in a buffer containing 300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 20 mM imidazole, 1% n-dodecyl-β-D-maltoside (DDM) and 0.1% cholesteryl hemisuccinate (CHS). Insoluble materials were removed by ultracentrifugation (138,000 g, 30 min, 4 °C), and the supernatant was mixed with 5 ml Ni-NTA Superflow (Qiagen). After binding for 1 h, the resin was washed with 300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 20 mM imidazole, 0.1% DDM and 0.01% CHS, and YidC was eluted with the same buffer supplemented with imidazole to a final concentration of 300 mM at 4 °C. The N-terminal residues and the His$_8$ tag or GFP-His$_8$ tag were cleaved by a His-tagged tobacco etch virus (TEV) protease, and the sample was reloaded onto the Ni-NTA column (5 ml) to remove the protease. The flow-through fraction containing YidC was collected, concentrated and further purified on a Superdex 200 10/300 GL column (GE Healthcare) in 300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 0.1% DDM and 0.01% CHS at 4 °C. For crystallization, the purified protein was concentrated to 6 mg ml^{-1} with a centrifugal filter device (Millipore, 50 kDa molecular weight cutoff) and dialysed against the following buffer (1 mM Tris-HCl, pH 8.0, 0.05% DDM and 0.005% CHS). For mercury derivatization, the Y150C YidC$_{27-266}$ mutant was purified and incubated with 2 mM methyl mercury chloride for 1 h before crystallization.

Crystallization and heavy-atom derivatization. The protein was mixed with monoolein in a 2:3 protein to lipid ratio (w/w), using the twin-syringe mixing method[27]. Aliquots (50 nl) of the protein lipidic cubic phase (LCP) mixture were spotted on a 96-well sandwich plate and overlaid with 800 nl precipitant solution using a mosquito LCP crystallization robot (TTP LabTech). The crystals of YidC$_{27-266}$ and Yid C$_{27-267}$ were grown at 20 °C in reservoir solutions containing 28–32% poly(ethyleneglycol) 500 dimethylether (PEG500DME), 2.5 mM CdCl$_2$ and 100 mM Na-(CH$_3$)$_2$AsO$_2$, pH 6.0; and 24–26% PEG500DME, 10 mM CuCl$_2$, 200 mM NH$_4$COOH and 100 mM MES-NaOH, pH 6.0, respectively. The heavy-atom derivatized crystals were obtained by co-crystallization of the Y150C mutant and methyl mercury chloride in the same reservoir solution used for the YidC$_{27-266}$ crystals. The crystals grew to full size in 2–3 weeks. The crystals were flash-cooled, using reservoir solution supplemented with 20% PEG500DME and 20% glycerol as a cryoprotectant and then stored in liquid nitrogen.

Data collection and structure determination. X-ray diffraction data sets were collected by the helical data collection method on beamline BL32XU at SPring-8, using a micro beam with a 1-μm width and a 10-μm height[28]. Diffraction data were processed using the program HKL-2000 (HKL Research) or the program XDS[29]. One Hg atom site was identified with the program SHELXD[30]. The initial phases were calculated using the program SHARP[31], followed by solvent flattening with SOLOMON[32]. The main chain was traced by automated model building using the program RESOLVE[34]. The model was further built manually using COOT[34] and refined using PHENIX[35]. The structures of YidC$_{27-266}$ and YidC$_{27-267}$ were determined by molecular replacement, using the program PHASER[36]. The Ramachandran plots were calculated with the program RAMPAGE[37]. Data collection and refinement statistics are provided in Extended Data Table 1. The YidC$_{27-266}$ crystal contains one molecule in the asymmetric unit. The YidC$_{27-267}$ crystal contains two molecules in the asymmetric unit (Mol A and Mol B), but these molecules do not form a face-to-face dimer. The figures depicting the molecular structures were prepared using CueMol (http://www.cuemol.org/).

Bacterial strains and plasmids for *in vivo* functional analysis. The *B. subtilis* strains, plasmids and DNA oligonucleotides used are listed in Supplementary Tables 1–3 and were constructed as described in the Supplementary Methods.

β-Galactosidase activity assay and immunoblotting. *B. subtilis* cells were cultured at 37 °C in LB or CH medium. Aliquots (500 μl) of cultures at an optical density at 600 nm of ~0.5 were harvested and used for β-galactosidase activity assays and immunoblotting. The β-galactosidase activities were measured as described previously[38,39]. For immunoblotting, a 500 μl aliquot of the culture was mixed with 56 μl 50% trichloroacetic acid and incubated on ice for at least 5 min. The cells were precipitated by centrifugation (4 °C, 15,000 r.p.m., 5 min), washed with 1 ml 1 M Tris-HCl, pH 8.0, resuspended in SB buffer (33 mM Tris-HCl, pH 8.0, 40% sucrose and 1 mM EDTA)

containing 1 mg ml^{-1} lysozyme (Sigma) and incubated at 37 °C for 10 min. The cells were then solubilized by adding an equal volume of 2× SDS loading buffer (4% SDS, 0.1 M Tris-HCl, pH 6.8, 30% glycerol and 10 mM dithiothreitol (DTT)) and subjected to immunoblotting, using either anti-FLAG (Sigma) or anti-SpoIIIJ antibodies, as described previously[40]. Antiserum production is described in the Supplementary Information.

Growth complementation assay. *B. subtilis* cells were cultured at 37 °C in LB medium containing 100 μg ml^{-1} spectinomycin and 1 mM IPTG. Aliquots (3.5 μl) of serially diluted (10^{-1}–10^{-5}), fully grown cultures were spotted on spectinomycin-containing LB agar plates with or without 1 mM IPTG and incubated at 37 °C for 15 h.

***In vivo* photo-crosslinking assay.** *B. halodurans yidC-FLAG*(Sigma-Aldrich)-*His$_8$* and *B. subtilis mifM* were cloned into the NcoI and BamHI sites of MCS1 and the NdeI and XhoI sites of MCS2 in pETDuet (Novagen), respectively. An amber mutation, TAG, was introduced into *yidC* by site-directed mutagenesis. *mifM*-deletion plasmids were generated by restriction digestion of the plasmids with SalI and XhoI and then ligation. *E. coli* BL21(DE3) cells harbouring two plasmids, pEVOL-pBpF (Addgene) and a pETDuet-based plasmid expressing YidC and MifM, were grown at 37 °C in M9-glucose medium supplemented with 1 mM pBpa[22] and appropriate antibiotics until mid-logarithmic phase, and then induced with 1 mM IPTG for 30 min. Portions of the culture (1 ml for isolation of total membrane and 2 ml for purification of YidC by Ni-NTA chromatography) were transferred to a dish and exposed to ultraviolet radiation (365 nm) for 5 min, by using a B-100AP ultraviolet lamp (UVP) at a distance of 5 cm. The irradiated cells were collected by centrifugation at 7,000 g for 1 min at 4 °C, suspended in 300 μl 10 mM Tris-HCl buffer, pH 8.0, containing 1 mM EDTA-Na and 0.1 mM 4-(2-aminoethyl)benzenesulphonyl fluoride hydrochloride, and disrupted by freeze-thawing and sonication (Qsonica) with cooling on ice. Cellular debris was separated by centrifugation at 9,000 g for 1 min at 4 °C and used for the isolation of the total membrane fraction or for the purification of YidC. The total membrane fraction was isolated by ultracentrifugation at 100,000 g for 20 min at 4 °C. For the purification of YidC, the membrane proteins were solubilized in SC buffer containing 20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 20 mM imidazole-HCl, pH 8.0, 1% DDM, 0.1% CHS and 0.1 mM 4-(2-aminoethyl) benzenesulphonyl fluoride hydrochloride for 30 min at 4 °C. Insoluble materials were removed by ultracentrifugation at 100,000 g for 20 min at 4 °C. The supernatant was mixed with 0.1 mL Ni-NTA agarose (Qiagen) that had been pre-equilibrated with SC buffer. After binding for 30 min at 4 °C, the resin was washed with SC buffer containing 0.1% DDM, and then YidC was eluted with SC buffer supplemented with 300 mM imidazole-HCl (pH 8.0). Proteins were separated by SDS–PAGE and detected by immunoblotting using anti-FLAG (Sigma) or anti-MifM[40] antibodies.

Molecular dynamics simulations. The simulation system included YidC, phosphoryl oleoyl phosphatidylethanolamine (POPE), water molecules and 150 mM NaCl. At first, the position and orientation of YidC in the POPE lipid bilayer were optimized using implicit solvent and membrane models (T.M. and Y. Sugita, manuscript in preparation). Next, the disordered region in the C2 loop (residues 200–213) was modelled, using the program MODELLER[41]. The missing atoms, including hydrogens in the protein, were built with the program VMD[42]. Finally, the periodic boundary system, including the explicit solvent and the POPE lipid bilayer[43], was prepared. The resultant size of the simulation box was 96 Å × 96 Å × 96 Å. The net charge of the solute was neutralized by adding chloride and sodium ions. The molecular topologies and parameters from CHARMM36 were used[43].

Molecular dynamics simulations were performed with the program NAMD (version 2.8)[44]. The system was first energy minimized for 1,000 steps with fixed positions for the non-hydrogen atoms, and then for another 1,000 steps with 10 kcal mol^{-1} restraints for the non-hydrogen atoms. Next, we performed a long equilibration run of 50 ns in the canonical (NVT) ensemble (300 K, 96 Å × 96 Å × 96 Å volume), with 10 kcal mol^{-1} restraints for protein non-hydrogen atoms and 0.1 kcal mol^{-1} restraints for water molecules, to optimize the locations of the lipid molecules around the protein, especially around the hydrophilic groove. Finally, equilibration was performed for 5 ns in the isothermal–isobaric (NPT) ensemble (300 K, 1.0 atm), with 10 and 0.1 kcal mol^{-1} restraints for the protein main chain and side chain atoms, respectively. The production process was performed for 1,000 ns. Constant temperature was maintained by using Langevin dynamics. Constant pressure was maintained by using the Langevin piston Nosé–Hoover method[45]. Long-range electrostatic interactions were calculated using the particle mesh Ewald method[46].

Determination of the molecular mass of YidC in detergent solution by using SEC-MALLS. The instrument set-up used for the SEC-MALLS experiment consisted of a Prominence HPLC system (Shimadzu) with an SPD-20A ultraviolet absorbance detector connected in series with a DAWN HELEOS II light-scattering detector (Wyatt Technology) and an Optilab T-rEX refractive index detector (Wyatt Technology). Analytical size exclusion chromatography was performed using a Superdex 200 10/300 column (GE Healthcare) equilibrated with buffer containing 300 mM NaCl, 20 mM Tris-HCl, pH 8.0, 0.1% DDM and 0.01% CHS. Purified YidC$_{27-266}$

(60 μg) was injected onto the column, and the elution was monitored in-line with three detectors. A 658.0-nm wavelength laser was used in the light scattering experiment. The data were corrected for the volume delay of ultraviolet radiation between the other detectors and were analysed using ASTRA software (Wyatt Technology). The molecular masses of the protein–micelle complex, the micelle and the protein were determined as described previously[47,48]. The refractive index increment (dn/dc) in which n is the refractive index and c is the concentration of the mixture of DDM and CHS in buffer, containing 300 mM NaCl and 20 mM Tris-HCl, pH 8.0, was determined offline using an Optilab T-rEX refractive index detector with a 658.0-nm wavelength laser, as described previously[49].

26. Nishizawa, T. *et al.* Structural basis for the counter-transport mechanism of a H$^+$/Ca^{2+} exchanger. *Science* **341,** 168–172 (2013).
27. Caffrey, M. Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Annu. Rev. Biophys.* **38,** 29–51 (2009).
28. Hirata, K. *et al.* Achievement of protein micro-crystallography at SPring-8 beamline BL32XU. *J. Phys. Conf. Ser.* **425,** 012002 (2013).
29. Kabsch, W. XDS. *Acta Crystallogr. D* **66,** 125–132 (2010).
30. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58,** 1772–1779 (2002).
31. de La Fortelle, E., Irwin, J. J. & Bricogne, G. SHARP: a maximum-likelihood heavy-atom parameter refinement program for the MIR and MAD methods. *Methods Enzymol.* **276,** 472–494 (1997).
32. Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **52,** 30–42 (1996).
33. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55,** 849–861 (1999).
34. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66,** 486–501 (2010).
35. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
36. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
37. Lovell, S. C. *et al.* Structure validation by Cα geometry: φ, ψ and Cβ deviation. *Proteins* **50,** 437–450 (2003).
38. Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory Press, 1972).
39. Rubio, A., Jiang, X. & Pogliano, K. Localization of translocation complex components in *Bacillus subtilis*: enrichment of the signal recognition particle receptor at early sporulation septa. *J. Bacteriol.* **187,** 5000–5002 (2005).
40. Chiba, S. *et al.* Recruitment of a species-specific translational arrest module to monitor different cellular processes. *Proc. Natl Acad. Sci. USA* **108,** 6073–6078 (2011).
41. Šali, A. *et al.* Evaluation of comparative protein modeling by MODELLER. *Proteins* **23,** 318–326 (1995).
42. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38 (1996).
43. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114,** 7830–7843 (2010).
44. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
45. Feller, S. E., Zhang, Y., Pastor, R. W. & Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *J. Chem. Phys.* **103,** 4613 (1995).
46. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089 (1993).
47. Hayashi, Y., Matsui, H. & Takagi, T. Membrane protein molecular weight determined by low-angle laser light-scattering photometry coupled with high-performance gel chromatography. *Methods Enzymol.* **172,** 514–528 (1989).
48. Slotboom, D. J., Duurkens, R. H., Olieman, K. & Erkens, G. B. Static light scattering to characterize membrane proteins in detergent solution. *Methods* **46,** 73–82 (2008).
49. Strop, P. & Brunger, A. T. Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci.* **14,** 2207–2211 (2005).
50. Chiba, S. & Ito, K. Multisite ribosomal stalling: a unique mode of regulatory nascent chain action revealed for MifM. *Mol. Cell* **47,** 863–872 (2012).
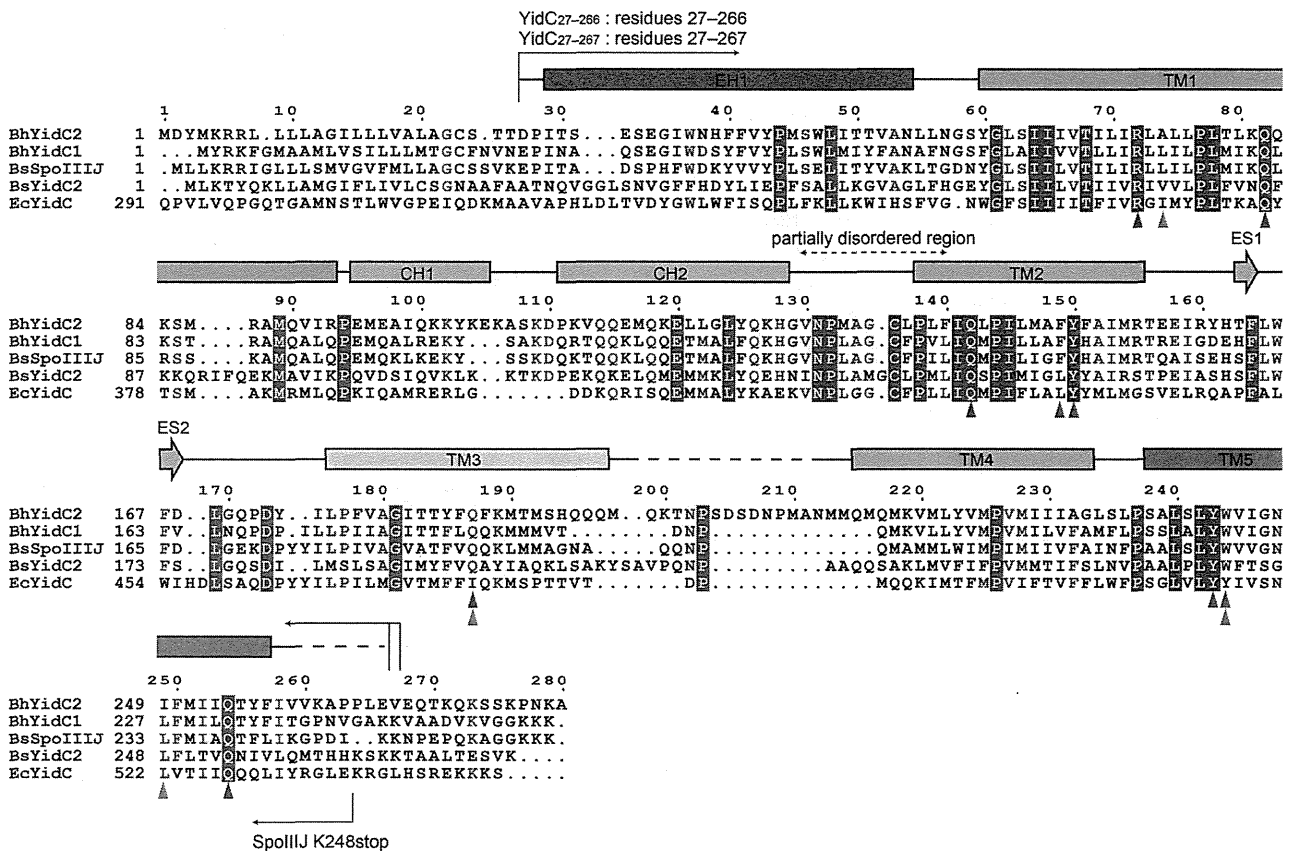
Extended Data Figure 1 | Multiple amino acid sequence alignment of YidC proteins. Sequence alignment of *Bacillus halodurans* YidC2 (BhYidC2), *B. halodurans* YidC1 (BhYidC1), *Bacillus subtilis* SpoIIIJ (BsSpoIIIJ), *B. subtilis* YidC2 (BsYidC2) and *Escherichia coli* YidC (EcYidC). The secondary structure of YidC$_{27-266}$ is indicated above the sequences. The α-helices (as described in the main text) and β-strands (ES1 and ES2 in the E2 region) are indicated by cylinders and arrows, resp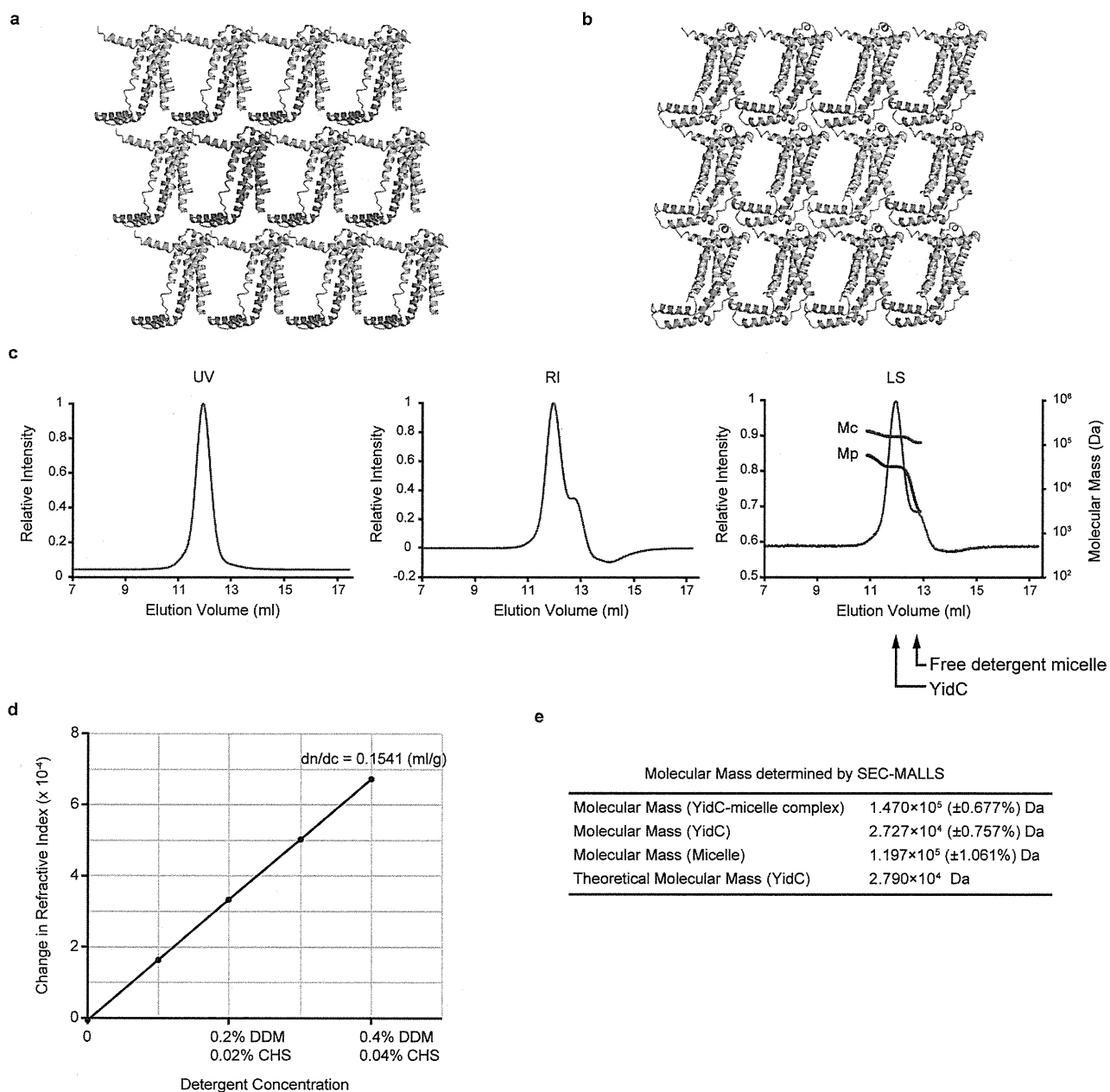ectively. Strictly conserved residues among the five molecules are highlighted in red boxes, and highly conserved residues are indicated by red letters. The hydrophilic and bulky residues that were mutated and the pBpa positions introduced into *B. halodurans* YidC2 are indicated by grey, green and blue triangles, respectively. The *spoIIIJ* K248stop derivative has a stop codon introduced at position 248, as indicated, and thereby expresses a SpoIIIJ mutant that lacks the C-terminal 14 residues.
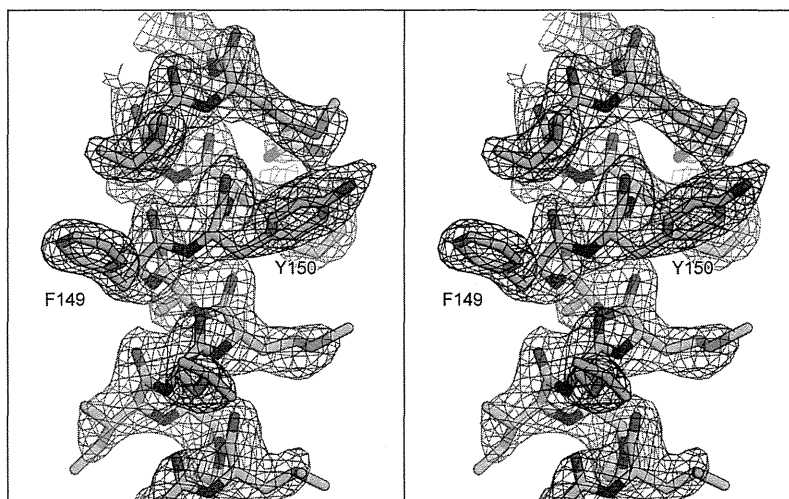
Extended Data Figure 2 | Electron density map of *B. halodurans* YidC. Stereo view of the $2mF_O - DF_C$ electron density map of the TM2 helix, contoured at $1.1\,\sigma$, where m is the figure of merit and D is the SIGMA-A weighting factor. F_O and F_C are the observed and the calculated structure factor amplitudes, respectively.

c

UV

RI

LS

Mc
Mp

Free detergent micelle
YidC

d

dn/dc = 0.1541 (ml/g)

0.2% DDM
0.02% CHS

0.4% DDM
0.04% CHS

Detergent Concentration

e

Molecular Mass determined by SEC-MALLS	
Molecular Mass (YidC-micelle complex)	1.470×10^5 (±0.677%) Da
Molecular Mass (YidC)	2.727×10^4 (±0.757%) Da
Molecular Mass (Micelle)	1.197×10^5 (±1.061%) Da
Theoretical Molecular Mass (YidC)	2.790×10^4 Da

Extended Data Figure 3 | Monomeric *B. halodurans* YidC. a, The crystal packing of $YidC_{27-266}$, viewed from the plane of the membrane. The molecule in the asymmetric unit is coloured red. **b,** The crystal packing of $YidC_{27-267}$, viewed from the plane of the membrane. Two molecules (Mol A in light pink and Mol B in light blue) are in the asymmetric unit. **c,** The chromatograms show the ultraviolet (UV), refractive index (RI) and light scattering (LS) detector readings. The volume delays of UV between the other detectors were corrected. The traces were normalized to the peak maxima. The green and blue lines in the LS chromatogram indicate the calculated molecular masses of the protein–detergent complex and the protein, respectively. **d,** The RI of the mixture was measured in response to five concentration steps. The refractive index increment (dn/dc) of the mixture of DDM and CHS was determined using linear regression of the RI versus the concentration. **e,** The molecular mass values determined by SEC-MALLS and calculated from the amino acid sequence.
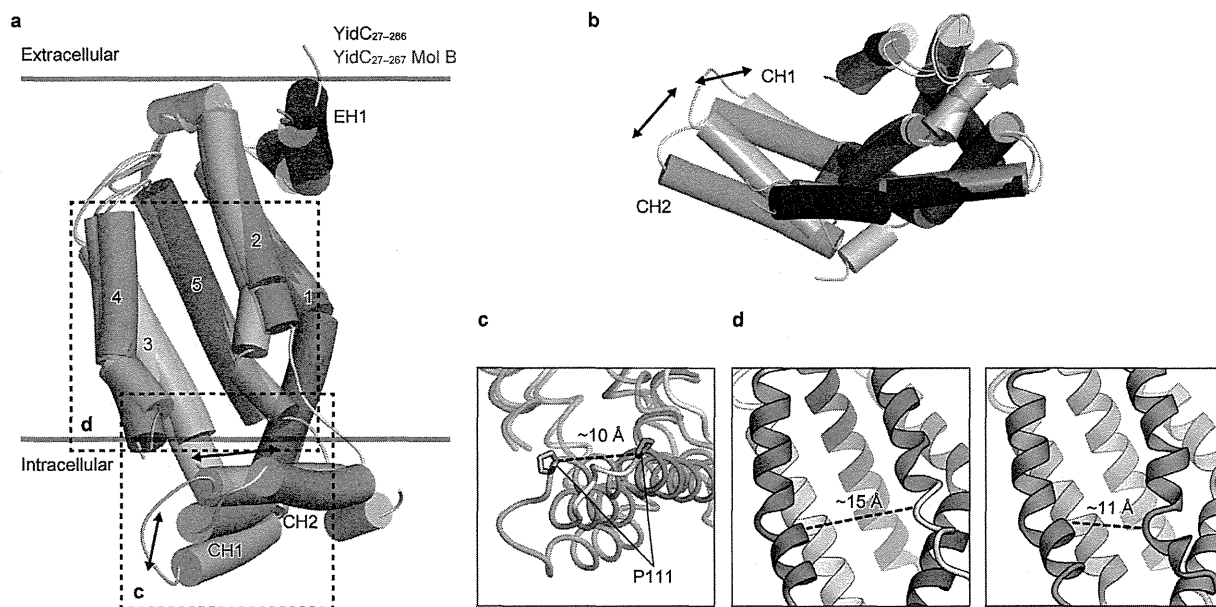
Extended Data Figure 4 | Structural flexibility of the hydrophilic groove and C1 region. a, b, Superimposition of the crystal structures of $YidC_{27-266}$ (coloured) and Mol B of $YidC_{27-267}$ (grey), viewed from the plane of the membrane (**a**) and from the extracellular space (**b**). The conformational changes observed in CH1 and CH2 are indicated by black arrows. **c,** Close-up view of the C1 region. The side chains of P111 are shown by stick models. In the $YidC_{27-267}$ structure, the arrangement of the C1 region with respect to the core region is rotated by ~35° compared with that in the $YidC_{27-266}$ structure. As a result, the tip of the C1 region is displaced by ~10 Å in the $YidC_{27-267}$ structure. **d,** Close-up views of the hydrophilic groove (left, $YidC_{27-266}$; right, Mol B of $YidC_{27-267}$). The distances between the Cα atoms of C136 and M221 are indicated by dashed lines.
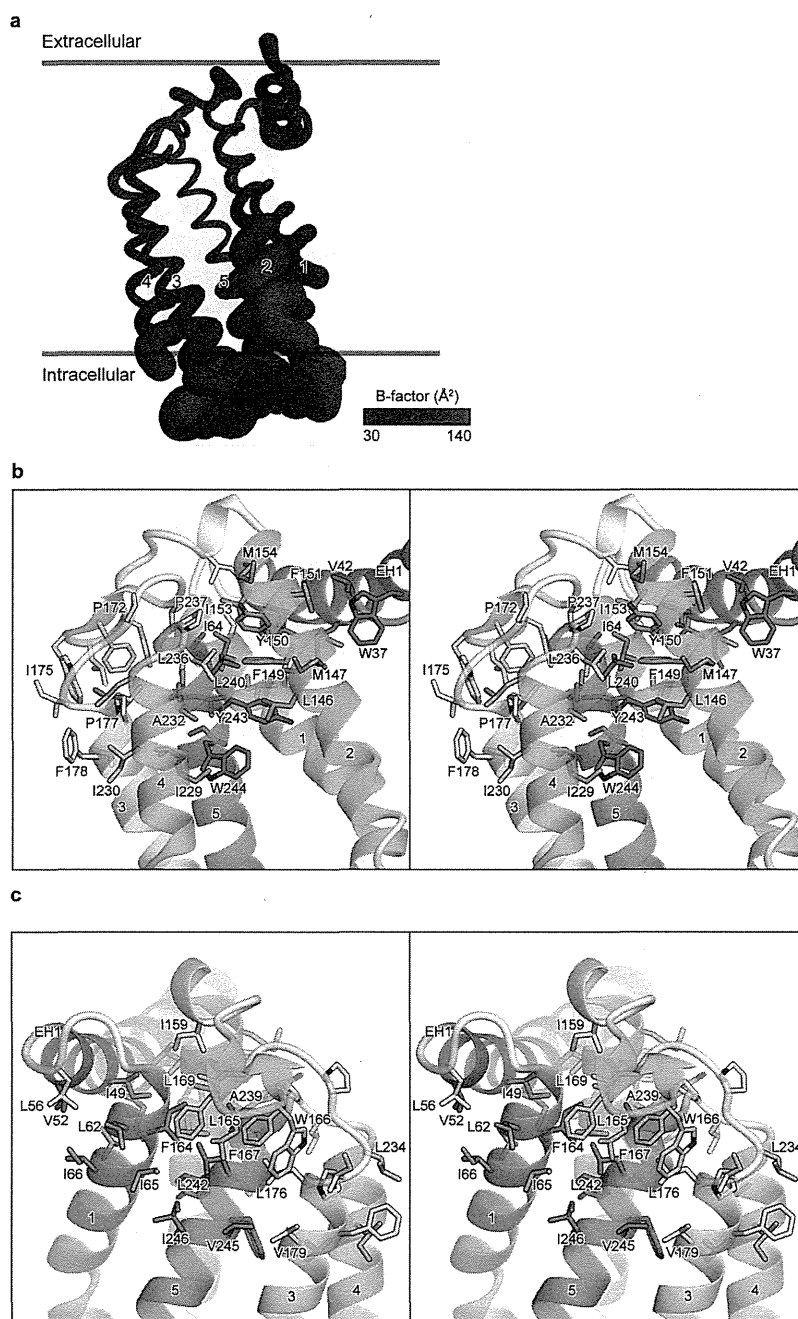
Extended Data Figure 5 | The hydrophobic core of *B. halodurans* YidC. **a,** The crystallographic B-factors are coloured in a gradient varying from blue to red, representing 30 to 140 Å². **b, c,** Stereo views of the hydrophobic core, showing the side chains of the hydrophobic residues.
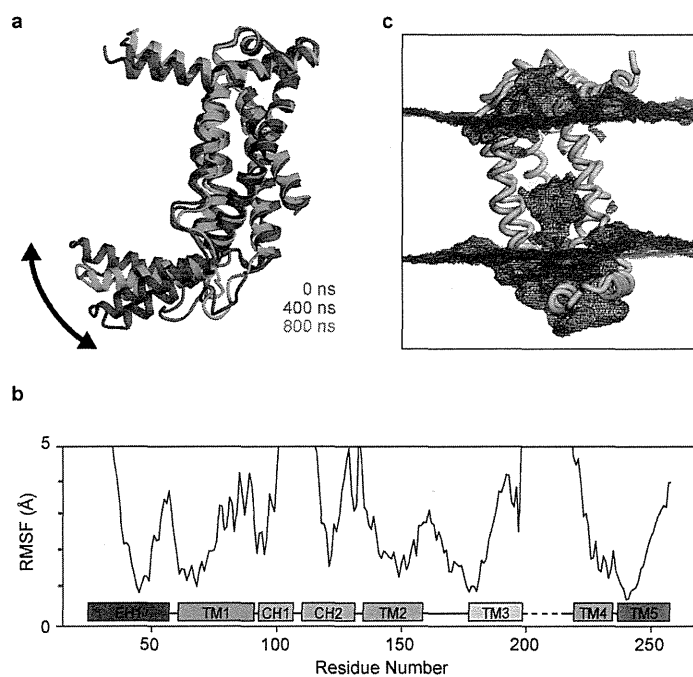
Extended Data Figure 6 | Molecular dynamics simulation of *B. halodurans* YidC for 1,000 ns in a lipid bilayer. a, Snapshots of the structure over the time course of the simulation at 400 ns intervals: 0 ns (blue), 400 ns (magenta) and 800 ns (light green). **b,** Root mean square fluctuation (r.m.s.f.) of YidC during the simulation. The secondary structure of YidC is indicated below the line. **c,** The water probability density map in the simulation, contoured at 0.001 molecules $\mathring{A}^{-3}$ ns^{-1}.