Site-Directed Alkylation Detected by In-Gel Fluorescence (SDAF) to Determine the Topology Map and Probe the Solvent Accessibility of Membrane Proteins

Yu-Hung Lin¹, Sung-Yao Lin¹, Guan-Syun Li¹, Shao-En Weng¹, Shu-Ling Tzeng², Yu-Hsuan Hsiao¹ & Nien-Jen Hu¹,³,⁴

The topology of helix-bundle membrane proteins provides low-resolution structural information with regard to the number and orientation of membrane-spanning helices, as well as the sidedness of intra/extra-cellular domains. In the past decades, several strategies have been developed to experimentally determine the topology of membrane proteins. However, generally, these methods are labour-intensive, time-consuming and difficult to implement for quantitative analysis. Here, we report a novel approach, site-directed alkylation detected by in-gel fluorescence (SDAF), which monitors the fluorescent band shift caused by alkylation of the EGFP-fused target membrane protein bearing one single introduced cysteine. In-gel fluorescence provides a unique readout of target membrane proteins with EGFP fusion from non-purified samples, revealing a distinct 5 kDa shift on SDS-PAGE gel due to conjugation with mPEG-MAL-5K. Using the structurally characterised bile acid transporter ASBT as an example, we demonstrate that SDAF generates a topology map consistent with the crystal structure. The efficiency of mPEG-MAL-5K modification at each introduced cysteine can easily be quantified and analysed, providing a useful tool for probing the solvent accessibility at a specific position of the target membrane protein.

The topology of helix-bundle membrane proteins summarizes two-dimensional information of the fold of the polypeptide chain across the cell membrane, i.e. the number, length and orientation of transmembrane segments. With increasing numbers of whole-genome sequences becoming available, in silico topology prediction, based on the hydrophobicity analysis of amino acid sequence and the “positive-in” law, is frequently applied to obtain a topological map of an integral membrane protein. While hydropathy plots illustrate the arrangement of transmembrane segments and thus provide a valuable guide for functional and structural studies, they constitute predictions and, therefore, are often erroneous and require experimental data for validation.

In the past decades, a number of methodologies have been developed for experimental validation of predicted topology maps (for detailed overviews see, for example). A common strategy is to construct a chimera fusion with reporter enzymes attached to the terminus of C-terminal truncated target membrane proteins, such as alkaline phosphatase PhoA, β-galactosidase LacZ, and β-lactamase bla. These enzymes show distinct activity in different subcellular compartments. Therefore, the enzyme activity in response to their cellular localization serves as a topological reporter. A similar concept was applied to probe protein glycosylation in eukaryotic cells. Because glycosylation takes place in the lumen of the endoplasmic reticulum (ER), only the glycosylation tag introduced in the luminal loop of ER membrane proteins can be glycosylated, probed by electrophoresis showing molecular

¹Graduate Institute of Biochemistry, National Chung Hsing University, 145 Xinda Rd., South Dist., Taichung City, 402, Taiwan, R.O.C. ²Institute of Medicine, Chung Shan Medical University, No.110, Sec. 1, Jianguo N. Rd., Taichung City, 40201, Taiwan, R.O.C. ³Rong Hsing Research Center for Translational Medicine, National Chung Hsing University, 145 Xinda Rd., South Dist., Taichung City, 402, Taiwan, R.O.C. ⁴Ph.D. Program in Transnational Medicine, National Chung Hsing University, 145 Xinda Rd., South Dist., Taichung City, 402, Taiwan, R.O.C. Correspondence and requests for materials should be addressed to N.-J.H. (email: njhu@nchu.edu.tw)
weight increase of oligosaccharides attachment\(^8,9\). Other strategies have been developed on the basis of specific interactions between the probe and target site that only occur in a solvent-exposed compartment. From the viewpoint of experiment design, proteolysis and epitope tags are used for proteolytic enzymes and monoclonal antibody recognition, respectively\(^10,11\). Because the probing enzymes and antibodies cannot cross the membrane, the distinct pattern obtained in such an analysis will show a lack of interactions between probe and target if the target loop is located intracellularly. Although the methods have successfully demonstrated their capability in topology determination, they share a common problem: it remains unclear whether the inserted tags, either enzymes or sites for glycosylation, proteolysis and immunogenic recognition would disturb the folding and expression of membrane proteins that are subjected to these analyses.

Substituted cysteine accessibility method (SCAM) originally developed by Karlin and co-workers is a versatile method using the covalent interaction between sulphydryl-specific reagents and cysteines for functional, structural and dynamical studies of membrane proteins\(^12,13\). Because cysteine is non-bulky in size and relatively hydrophobic, cysteine substitution in most positions of a membrane protein is probably tolerable\(^1\). SCAM was then applied to define transmembrane orientation by Bagdanov and co-workers, denoted by SCAM\(^\text{TM}\), using membrane-impermeable sulphydryl-reactive reagents to probe the membrane-embedded and loop regions of membrane proteins in whole cells, right-side-out vesicles, inside-out vesicles, reconstituted proteoliposomes and intact organelles\(^3,14\). Based on the same concept, Kaback and co-workers developed the site-directed alkylation (SDA) strategies, by labelling the introduced cysteine residues of LacY with sulphydryl-reactive fluorophores and isotopes, and demonstrated the substrate-induced conformational change responsible for shifting the equilibrium toward the outward-facing state\(^15,16\). In spite of the effectiveness of the aforementioned methods, detection of the specific modification on target membrane proteins requires labels that afford unique signals for detection, for example by means of chemiluminescence, autoradiography or fluorescence, etc. In most cases, the target membrane proteins have to be purified after labelling.

Here, we present a novel SDA method, termed SDAF, for the determination of transmembrane topology and study of conformational dynamics of membrane proteins with a C-terminal EGFP fusion allowing specific detection of cysteine labelling without any further imaging treatment and protein purification. The sulphydryl-reactive reagent, methoxypolyethylene glycol maleimide 5 K (mPEG-MAL-5K), covalently attaches to solvent accessible cysteine residues on the surface of intact cells, resulting in a 5 kDa band shift observable by in-gel fluorescence. As illustrated in Fig. 1, being membrane impermeable, mPEG-MAL-5K can only react with cysteine residues in the extracellular loops when performing whole cell PEGylation; it cannot gain the access to cysteine residues in the intracellular loops without disrupting the cell membrane. Application of the SDAF methodology is demonstrated by mapping the topology of the apical Na\(^+/\) bile acid transporter homolog from Neisseria meningitidis (ASBT\(_{\text{NM}}\)), the results of which were in agreement with the published crystal structure\(^17\). Moreover, the degree of PEGylation at given position can be evaluated densitometrically by the ratio of the fluorescence emission of shifted and non-shifted bands. For the example presented, the PEGylation levels of each selected cysteine substitution were in excellent agreement with the solvent accessibility of those residues in ASBT\(_{\text{NM}}\). In summary, SDAF provides
an efficient and direct method for topological mapping and characterization of the conformational variations of membrane proteins in native membrane environment.

Results

**mPEG-MAL-5K gains access to the cysteine residue on the intracellular side after cell membrane disruption by sonication.** Because overexpressed ASBTNM is localized at the inner membrane of *E. coli* cells, we first characterize whether mPEG-MAL-5K can permeate through the outer membrane. The outer membrane of *E. coli* is porous allowing passage of small molecules with the range of 500–700 Da. Nevertheless, it has been shown that mPEG-MAL-5K can label a solvent accessible cysteine residues of an inner membrane protein on intact *E. coli* cell membrane. To validate the hypothesis, we constructed a cysteine-free pWaldo-cfASBTNM-EGFP-His8 (see Methods) and introduced a single cysteine mutation on the extracellular loop for whole cell PEGylation assay. We initially substituted cfASBTNM A275 (Supplementary Fig. S1a), which is located at the C-terminus of TM9b exposed to the periplasm, with cysteine. The side chain of residue 275 is at a very solvent accessible position and the cysteine sulfhydryl group is situated at the entrance of the outward-facing vestibule for substrate uptake. The PEGylation results showed that cfASBTNM A275C was chemically modified with mPEG-MAL-5K as the fluorescent protein band shifted approximately 5 kDa (Figs 1c, 2a and Supplementary Fig. S1b), indicating mPEG-MAL-5K can permeate the outer membrane of *E. coli*. The labelling efficiency, as exemplified in Supplementary Fig. S2 and calculated from Eq. (1), was dose-dependent and reached a saturation (approximately 80%) at a molar ratio of 1:70,000 (ASBTNM: mPEG-MAL-5K; Fig. 2b and Supplementary Fig. S1b). As the overexpressed target protein can be quantified using whole cell fluorescence counts, it is straightforward to perform the assay with a constant protein to reagent molar ratio.

It is noted that a band below the fusion protein standing for the size of free EGFP was observed, probably due to non-specific proteolysis during cell lysis in SDS-PAGE sample preparation using *E. coli* whole cells. EGFP...
contains two native cysteine residues, C48 and C70. In a properly folded EGFP molecule, these two cysteines are at the reduced form because they are too distant to form disulphide bond (24 Å). C70 is deeply buried in the β-barrel and C48 is relatively more accessible for sulphydryl reagents. At the molar ratio of 1:90,000, a faint band shift above the free EGFP is becoming visible (Supplementary Fig. S1b), indicating that EGFP can be labelled with mPEG-MAL-5K in the presence of extremely excessive sulphydryl reagent. Consequently, we chose the molar ratio of 1:70,000 to perform the subsequent whole cell PEGylation assays.

In order to probe the intracellular loops, we introduced a cysteine at D61 located in the intracellular loop connecting helices TM2 and 3 (Supplementary Fig. S3a). D61C showed no band shifting after PEGylation using intact cells (Figs 1e, 2c and Supplementary Fig. S3b, Control), suggesting mPEG-MAL-5K cannot permeate through the inner membrane. However, the same experiment using disrupted cells revealed a significant band shift with 5 kDa (Figs 1f, 2c and Supplementary Fig. S3b). The labelling efficiency using disrupted membranes also demonstrated a dose-dependent increase and reached a saturation at a molar ratio of 1:5,000 (ASBT

Examining the mPEG-MAL-5K accessibility of the two native cysteine residues in EGFP. In the pWaldo-cfASBT

Topological mapping of ASBT

PEGylation profiles of ASBT

pathway for individual cysteine substitution (Fig. 4a). PEGylation treatment was performed using intact *E. coli* cells overexpressing the mutants (Supplementary Fig. S5b). The uptake activities of these mutants were assessed, revealing a significant variation (Supplementary Fig. S8b). These mutants, especially G267C, located on the putative substrate binding site on the ASBTNM core domain midway across the membrane, revealed ~20% uptake activity, probably due to its important role in substrate binding. As shown in the in-gel fluorescence images, the band shift of the mutants (A275C and I40C) positioned on the extracellular side was the most prominent (Fig. 4b and Supplementary Fig. S7). PEGylation efficiency decreased as the position of the cysteine replacement proceeds toward the intracellular side. The PEGylation efficiency at each position was calculated and plotted as a PEGylation profile (Fig. 4c). As the replacements positioned below I47C (TM2) on the panel domain and V116C (TM4b) on the core domain showed extremely low PEGylation efficiencies, these positions (I47C, M48C, M51C, V116C and Y119C) can be inferred as not being solvent accessible from the extracellular side of intact cells. Based on these results, one can conclude that the lower regions of TM2 and TM4b are packed tightly, consistent with the outward-facing crystal structure of ASBTYf-E254A19. Furthermore, the data also corroborate the principle of alternating access mechanism of transporters: the substrate binding site in the core is exposed to either side of the membrane in an alternating fashion, but never simultaneously. In summary, the quantitative analysis delivering a PEGylation profile provides a useful method to evaluate the solvent accessibility of the substrate permeation pathway in native cell membrane environment.

**Discussion**

Previously reported cysteine-scanning mutagenesis methods provided useful tools to obtain low-resolution structural information of membrane proteins, including topology maps, helix bundle packing and characterization of the substrate-binding site. However, the time-consuming and labour-intense protocols required by these methods make them less practical for quantification of the amount of target proteins and the engineered cysteine modifications. SDAF constitutes a substantial advancement of such methods and addresses the current shortcomings without the need for Western blotting or protein purification for visualization of target membrane proteins. Instead, this new methodology employs site-directed alkylation with mPEG-MAL-5K and can be performed using *E. coli* whole cells overexpressing the target protein fused with EGFP21,23. As a result, the PEGylated and non-PEGylated target membrane proteins can be distinctly detected by SDS-PAGE gel using in-gel fluorescence, rendering the EGFP fusion protein in its folded state and thus capable of emitting fluorescence during/after the

![Figure 3. SDAF reveals a topological map of ASBTNM in accordance with the crystal structure. Top panel: whole cell in-gel fluorescence images of the cysteine mutants with indicated treatments. For NEM (+) experiments, the samples were treated with NEM before mPEG-MAL-5K. Middle panel: the topology map of ASBTNM determined by the crystal structure. Red and blue helices represent the components of panel and core domains, respectively. Residues with numbers represent the sites where cysteine residues were introduced by single point mutagenesis. A29, E153, N93, E220 and A279 are located in the extracellular loops; N2, S186, D61, N124, T247 and A309 are located in the intracellular loops. Lower panel: in-gel fluorescence images of the cysteine mutants with indicated treatments. Membrane permeabilization is achieved by sonication (+), where the *E. coli* cells were disrupted before the following treatments. See Methods for details.](image-url)
electrophoresis. The assay itself can be done in ~3 hrs, therefore significantly improving the throughput of panels of cysteine mutants. Importantly, the molar ratio of target and reagent can be easily controlled by analysing the whole cell EGFP fluorescence counts. Therefore, the mPEG-MAL-5K labelling efficiencies of each cysteine mutant are pre-normalized and can be quantitatively compared. Carrying out such quantitative analysis in previously published methods has been proven very challenging, thus rendering quantitative analysis with those protocols inconsistent.

As mPEG-MAL-5K cannot permeate the inner membrane, the cysteine residues introduced in the intracellular loop showed no band shift using E. coli whole cells for PEGylation. Cell lysis using sonication resulted in gaining access of mPEG-MAL-5K to the intracellular cysteine residues. We did not use detergents to lyse cell membrane for two reasons: (1) some detergents may disrupt the native folding of membrane proteins, and (2) detergent micelles may introduce steric hindrance preventing the contact of the alkylation reagent with the cysteine sulfhydryl group. Potentially, inside-out vesicles (ISOVs) could be employed to allow access to the cysteine residues residing in the intracellular loops; however, the preparation of ISOVs requires further characterization to confirm the formation of uniformly oriented vesicles.

As folded EGFP has two endogenous cysteines, C48 and C70 at the reduced form, an accurate control of the concentration of the sulphydryl reagent is required while performing SDAF experiments. We have shown that mPEG-MAL-5K is not permeable to the inner membrane of intact E. coli cells (Fig. 2c and Supplementary Fig. S3b). However, in the presence of extremely high concentration of mPEG-MAL-5K, such as the condition at 1:90,000 protein to mPEG-MAL-5K molar ratio, a band shift above the band of free EGFP is becoming visible (Supplementary Fig. S1b), suggesting excessive mPEG-MAL-5K may attach to the two endogenous cysteines in EGFP, although their accessibility to sulfhydryl reagents is limited in properly folded state20. While performing SDAF experiments using disrupted membranes, the concentration of mPEG-MAL-5K is added at much lower level because the membrane barrier is impaired and the cysteines on EGFP are more exposed to the sulfhydryl reagent. In this study, we used the protein to mPEG-MAL-5K ratios of 1:70,000 and 1:5,000 for whole cells and sonicator-disrupted membranes, respectively. To avoid the uncertainty in data interpretation, one has to perform extensive control experiments to determine the optimal molar ratio.

Although the use of C-terminally fused EGFP as the reporter has proven to be of great value in screening expression conditions and detergents for integral membrane proteins, it imposes a fundamental constraint: the target protein must have an intracellular C-terminus (Cin topology), otherwise the fused EGFP is not fluorescent. Extracellularly located EGFP has lost its ability to fluoresce due to being secreted in the unfolded state via the Sec system. Arriving in the oxidizing environment of the periplasm as an unfolded polypeptide chain, the two cysteine residues, C48 and C70, form intermolecular disulphide bonds which prevents folding of the β-barrel and formation of the fluorophore. A prediction of topologies of membrane proteins from 29 whole genome sequences demonstrated that 35% of multi-spanning membrane proteins have Cout topology, which, at present, are not amenable to the technique introduced here. To overcome this problem, a method could be employed that allows to convert the topology of membrane proteins from Cout to Cin by fusing glycophorin A (GpA), a single
membrane-spanning protein, on the C-terminus of the target protein. The authors of that study also demonstrated that, using GpA as the fusion, the C-terminus is redirected into the cytoplasm, allowing the downstream GFP to become fluorescent, and the functionality was not impaired by the large fusion tag. Alternatively, an EGFP variant called ‘superfolder GFP’ (sfGFP) could be used instead of EGFP; sfGFP revealed faster folding kinetics and remained fluorescent in oxidizing environments. It is also possible to utilize a cysteine-free GFP (cfGFP) variant that has been shown to possess comparable fluorescent brightness. In future studies, replacing the EGFP gene with sfGFP or cfGFP in the pWaldo vector may enable membrane proteins with C-terminal topology to be fluorescent for SDAF assays.

In the present study, we introduced SDAF as a convenient tool for mapping the topology of membrane proteins and showed that the PEGylation levels of cysteine mutants can be easily quantified using SDAF, and the levels are indicative of the solvent accessibility of the given location in the protein. Therefore, SDAF can be applied to study conformational changes of membrane proteins in native cell membranes. Kaback and coworkers have developed a series of alkylation methods using radioactive or fluorescent sulphydryl-reactive reagents (isotope-labelled NEM and TMRM), and demonstrated that ligand binding increases alkylation reactivity of cysteine replacements on the periplasmic side of LacY; i.e., LacY shifts to the outward-facing conformation in the presence of ligand. We also showed that PEGylation profiles obtained using the SDAF methodology provides a sensitive means to portray the substrate permeation pathway and may be applied to study substrate-induced dynamics of transporters. This can be achieved by comparing the PEGylation profiles using intact cells as well as permeabilized membranes, in the presence and absence of ligands and coupled ions. The systematic analysis based on data obtained with the SDAF methodology thus facilitates an understanding of alternating access mechanism for secondary active transporters.

In summary, SDAF is a versatile method allowing an efficient topology determination of multi-spanning membrane proteins, with α-helical or β-barrel folds, in a native membrane environment. The experimental data can be used to verify the in silico topology prediction of membrane proteins without known atomic structure. Additionally, the PEGylation profiles generated by SDAF can also be utilized in the studies of conformational dynamics of either soluble and membrane proteins.

Methods

Plasmid construction. The expression plasmid pWaldo-ASBTNM-EGFP-His8 containing the DNA fragment encoding ASBTNM was constructed as previously described. The two native cysteine residues of ASBTNM Cys107 and Cys108, were substituted with serine using QuickChange Lightning Site-Directed Mutagenesis kit (Agilent). The cysteine-free mutant protein, termed pWaldo-cfASBTNM-EGFP-His8, was used as template for single cysteine mutagenesis at positions A29, N93, E153, E220 and A279 residing in the extracellular loops, and N2, D61, N124, S186, T247 and A309 located in the intracellular loops. For evaluation of the solvent accessibility of the substrate permeation pathway constituted by helix TM2 of the panel domain and the discontinuous helices TM4b and TM9b of the core domain, we also systematically introduced individual cysteine residues at positions I40, P41, L44, I47, M48 and M51 (helix TM2, panel domain), T112, V116 and Y119 (helix TM4b, core domain), and N265, G267, A271 and A275 (helix TM4b, core domain), of which the side chains are facing the substrate permeation pathway (Fig. 4a). Each mutant was confirmed by DNA sequencing of the nucleotide sequence encoding ASBTNM.

Expression of cfASBTNM-EGFP and mutants. Escherichia coli CA3(DE3) transformed with pWaldo containing cfASBTNM with one single point mutation was grown in 10 mL LB broth in the presence of 50 μg/mL kanamycin at 37 °C. The overnight cultures were diluted 50-fold in 10 mL LB broth and ASBTNM mutants were induced with 0.4 mM IPTG when the OD600 reached 0.4, followed by overnight growth at 25 °C. 1 mL of overnight culture was subjected to centrifugation at 13,000 rpm using a benchtop centrifuge. The pellet was resuspended in 100 μL 1 × PBS and transferred to a 96-well microplate (Garnier) for fluorescence intensity measurement (λex = 485 nm, λem = 512 nm) using a spectrophotometer (TECAN). The fluorescence counts were converted to EGFP concentration using an in-house standard curve.

Labelling with mPEG-MAL-5K. Approximately 1.4 μg EGFP worth of culture was taken from the rest of the overnight culture for mPEG-MAL-5K labelling. As the molecular weights of EGFP and cfASBTNM in the pWaldo-ASBTNM-EGFP-His8 construct are 27 kDa and 32 kDa, respectively, 1.4 μg of EGFP corresponded to 1.6 μg of ASBTNM in the fusion construct. The aliquoted culture was subjected to centrifugation and washed two times using PEGylation buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10% (v/v) glycerol). The pellet was resuspended in 350 μL PEGylation buffer and methoxypolyethylene glycol maleimide 5,000 (mPEG-MAL-5K, Sigma-Aldrich) was added at a final concentration of 10 mM. Hence, the molar ratio of ASBTNM to mPEG-MAL-5K was approximately 1:70,000. The mixture was incubated at 27 °C for 1 h in the dark. 20 mM β-ME was added to terminate the reaction, followed by centrifugation at 13,000 rpm for 1 min. For control experiments, free cysteine residues were blocked by adding NEM at a molar ratio of 1:1 before mPEG-MAL-5K treatment.

Disruption of cell membranes. For the mutants carrying engineered cysteine residues in the intracellular loops of ASBTNM (N2, D61c, N124C, S186C, T247C and A309C), the E. coli cell membrane was permeabilized using sonication. 60 mL of IPTG-induced culture of each ASBTNM mutant was grown at 25 °C overnight and harvested by centrifugation as described above. The cell pellets were resuspended in 5 mL sonication buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 15 mM EDTA and 100 μM PMSE). The cell membranes were disrupted on ice using a probe sonicator (MISONIX XL-2020) at a power amplitude of 2.5 for 4 cycles of 45 s bursts within 60 s intervals. The remaining intact cells were removed by low speed centrifugation at 6,000 rpm for 10 min using a benchtop centrifuge. The supernatant was collected and the crude membranes were further collected using a
benchtop ultracentrifuge at 43,000 rpm for 1 hr. The membrane pellet was resuspended in 5 mL PEGylation buffer using a Dounce glass homogenizer. The fluorescence count of the crude membranes was measured as mentioned above and the equivalent of 1.6 μg ASBT<sub>NM</sub> of crude membranes was aliquoted. mPEG–MAL–5K was added to the crude membranes at a molar ratio of approximately 1:5,000 (ASBT<sub>NM</sub>·mPEG–MAL–5K). As a control experiment, free cysteine residues were blocked by adding NEM at a molar ratio of 1:1 before mPEG–MAL–5K treatment.

**Imaging using in-gel fluorescence.** In-gel fluorescence using whole cells was performed as described previously<sup>21</sup>. Briefly, the cell pellet was resuspended in 10 μL 1 × PBS and 10 μL 2 × sample loading buffer (200 mM Tris–HCl pH 8.8, 20% (v/v) glycerol, 5 mM EDTA pH 8.8, 0.02% Bromophenol Blue, and 4% SDS). 0.3 μL Benzonase Nuclease (Sigma) was added to digest the exposed chromosomal DNA. The mixture was subjected to centrifugation at 15,000 rpm, the supernatant was loaded onto a 4–20% Tris-Glycine SDS-PAGE gel and electrophoresis was performed on ice. For in-gel fluorescence of disrupted cell membranes, 10 μL of PEGylated crude membranes were mixed with 10 μL sample buffer. The mixture was subjected to SDS-PAGE electrophoresis performed on ice. To detect the fluorescent band of cASBT<sub>NM</sub>–EGFP fusion proteins, the SDS-PAGE gel was analysed densitometrically using the Gel Doc™ EZ imaging system (Bio-Rad) and a blue tray. Densitometric analysis of the fluorescent bands was carried out using the software Image Lab (Bio-Rad). The labelling efficiency of a particular cysteine residue was calculated by dividing the fluorescence density of the shifted band by the total density of the shifted and non-shifted bands of ASBT<sub>NM</sub>–EGFP.

![Equation](https://doi.org/10.1038/s41598-019-49292-w)

**Substrate uptake activity assay.** Substrate uptake activity of ASBT<sub>NM</sub> mutants was conducted in *E. coli* C43(DE3) whole cells according to the previously published method<sup>37</sup>. Briefly, 1.5 mL LB culture of *E. coli* cells overexpressing the WT ASBT<sub>NM</sub>–EGFP and cysteine mutants were harvested and the amount of target protein was normalised based on the EGFP fluorescence counts. The cell pellets were resuspended in 150 μL uptake buffer consisting of 100 mM NaCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 10 mM Tris–HCl pH:7.5. Cells were incubated at 37°C with uptake buffer containing 4 μM taurocholate supplemented with 0.17 μM [3H(G)]-taurocholate (15.4 Ci mmol⁻¹; Perkin Elmer) for 10 mins. Substrate uptake was terminated by adding 150 μL ice-cold stop buffer containing 100 mM NaCl, 1 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 10 mM Tris–HCl pH:7.5. 1 mM Taurocholic Acid immediately followed by vacuum filtration (QIAvac 24 Plus Vacuum Manifold, Qiagen) and four cycles of wash using 5 mL uptake buffer. The radioactivity of the internalized [3H(G)]-taurocholate was measured using scintillation counting. Each uptake experiment was performed in triplicate. The basal uptake was measured in triplicate by using C43(DE3) *E. coli* transformed with pWaldo–EGFP. Specific uptake was obtained by the subtraction of basal uptake from total uptake.

**References**

1. Nilsson, J., Persson, B. & von Heijne, G. Consensus predictions of membrane protein topology. *FEBS Lett* **486**, 267–269 (2000).
2. van Klompenburg, W., Nilsson, L., von Heijne, G. & de Kruijff, B. Anionic phospholipids are determinants of membrane protein topology. *The EMBO journal* **16**, 4261–4266 (1997).
3. Bogdanov, M., Zhang, W., Xie, J. & Dowlan, W. Transmembrane protein topology mapping by the substituted cysteine accessibility method (SCAM(TM)): application to lipid-specific membrane protein topogenesis. *Methods* **36**, 148–171, https://doi.org/10.1016/j.meth.2004.11.002 (2005).
4. van Geest, M. & Lolkema, J. S. Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiology and molecular biology reviews: MMBR* **64**, 13–33 (2000).
5. Manoil, C. & Beckwith, J. A genetic approach to analyzing membrane protein topology. *Science* **233**, 1403–1408 (1986).
6. Froshauer, S., Green, G. N., Boyd, D., McGovern, K. & Beckwith, J. Genetic analysis of the membrane insertion and topology of MalF, a cytoplasmic membrane protein of *Escherichia coli*. *J Mol Biol* **200**, 501–511 (1988).
7. Broome-Smith, J. K., Tadayyon, M. & Zhang, Y. Beta-lactamase as a probe of membrane protein assembly and protein export. *Molecular microbiology* **4**, 1637–1644 (1990).
8. Chang, X. B., Hou, Y. X., Jensen, T. J. & Riordan, J. R. Mapping of cystic fibrosis transmembrane conductance regulator membrane topology by glycosylation site insertion. *FEBS Lett* **364**, 34190–34195 (1998).
9. Hresko, R. C., Kruse, M., Strube, M. & Mueckler, M. Topology of the Glut 1 glucose transporter deduced from glycosylation site insertion. *J Biol Chem* **269**, 18572–18575 (1994).
10. Shimon, M. B., Goldshleger, R. & Karlish, S. J. Specific Cu<sup>2+</sup>-catalyzed oxidative cleavage of Na,K-ATPase at the extracellular surface. *J Biol Chem* **273**, 34190–34195 (1998).
11. Toyoda, T. et al. Antibody-scanning and epitope-tagging methods: molecular mapping of proteins using antibodies. *Curr Protein Pept Sci* **1**, 303–308 (2000).
12. Akbas, M. B., Stauffer, D. A., Xu, M. & Karlin, A. Acetylcholine receptor channel structure probed in cysteine-substitution mutants. *Science* **288**, 307–310 (1992).
13. Karlin, A. & Akbas, M. H. Substituted-cysteine accessibility method. *Methods in enzymology* **293**, 123–145 (1998).
14. Bogdanov, M. Mapping of Membrane Protein Topology by Substituted Cysteine Accessibility Method (SCAM). *Methods Mol Biol* **1615**, 105–128, https://doi.org/10.1007/978-1-4939-7033-9_9 (2017).
15. Kaback, H. R. et al. Site-directed alkylation and the alternating access model for LacY. *Proc Natl Acad Sci USA* **104**, 491–494, https://doi.org/10.1073/pnas.0609968104 (2007).
16. Nie, Y., Ersmolova, N. & Kaback, H. R. Site-directed alkylation of LacY: effect of the proton electrochemical gradient. *J Mol Biol* **374**, 356–364, https://doi.org/10.1016/j.jmb.2007.09.006 (2007).
17. Hu, N. J., Iwata, S., Cameron, A. D. & Drew, D. Crystal structure of a bacterial homologue of the bile acid sodium symporter ASBT. *Nature* **478**, 408–411, https://doi.org/10.1038/nature10450 (2011).
18. Yang, N. J. & Hinner, M. J. Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol Biol* **1266**, 29–53, https://doi.org/10.1007/978-1-4939-2272-7_3 (2015).
19. Zhou, X. et al. Structural basis of the alternating-access mechanism in a bile acid transporter. *Nature* **505**, 569–573, https://doi.org/10.1038/nature12811 (2014).
20. Ormo, M. et al. Crystal structure of the Aequorea victoria green fluorescent protein. Science 273, 1392–1395, https://doi.org/10.1126/science.273.5280.1392 (1996).

21. Drew, D. E., von Heijne, G., Nordlund, P. & de Gier, J. W. Green fluorescent protein as an indicator to monitor membrane protein overexpression in Escherichia coli. FEBS Lett 507, 220–224 (2001).

22. Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. Rapid protein-folding assay using green fluorescent protein. Nat Biotechnol 17, 691–693, https://doi.org/10.1038/10904 (1999).

23. Drew, D., Lerch, M., Kunji, E., Slotboom, D. J. & de Gier, J. W. Optimization of membrane protein overexpression and purification using GFP fusions. Nat Methods 3, 303–313, https://doi.org/10.1038/nmeth0406-303 (2006).

24. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305, 567–580, https://doi.org/10.1006/jmbi.2000.4315 (2001).

25. Hsieh, J. M. et al. Bridging the gap: a GFP-based strategy for overexpression and purification of membrane proteins with intra and extracellular C-termini. Protein science: a publication of the Protein Society 19, 868–880, https://doi.org/10.1002/pro.365 (2010).

26. Azim, D. E., Costantini, L. M. & Snapp, E. L. Superfolder, GFP is fluorescent in oxidizing environments when targeted via the Sec translocon. Traffic 12, 543–548, https://doi.org/10.1111/j.1600-0854.2011.01168.x (2011).

27. Suzuki, T. et al. Development of cysteine-free fluorescent proteins for the oxidative environment. PloS one 7, e37551, https://doi.org/10.1371/journal.pone.0037551 (2012).

28. Guan, L. & Kaback, H. R. Site-directed alkylation of cysteine to test solvent accessibility of membrane proteins. Nature protocols 2, 2012–2017, https://doi.org/10.1038/nprot.2007.275 (2007).

29. Lin, S.-Y. et al. Fluorophore Absorption Size Exclusion Chromatography (FA-SEC): An Alternative Method for High-Throughput Detergent Screening of Membrane Proteins. PloS one 11, e0157923, https://doi.org/10.1371/journal.pone.0157923 (2016).

Acknowledgements
We thank Dr. David Drew for providing the expression vector, and Dr. Andreas Hofmann for critical comments and diligent proofreading. The research work was supported by the Ministry of Science and Technology, Taiwan, R.O.C. (MOST 103-2311-B-005-007-MY3 and MOST 106-2311-B-005-013) and the joint grant of National Chung Hsing University and Chung Shan Medical University (NCHU-CSMU 10704).

Author Contributions
Y.-H.L. and N.-J.H. designed the experimental protocols of SDAF for intact E. coli whole cells. Y.-H.L. and N.-J.H. modified the protocols for permeabilized membranes. S.-Y.L., G.-S.L. and S.-E.W. performed the site-directed mutagenesis of ASBTNM and contributed to preparing Fig. 4. Y.-H.L. completed the data collection for topological mapping and contributed to preparing Figs 2 and 3. S.-L.T. reviewed the manuscript and provided advice on biochemistry aspects of the study. Y.-H.L. and N.-J.H. assisted in editing and preparing the figures. N.-J.H. wrote and edited the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-49292-w.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2019