Substrate-induced assembly and functional mechanism of the bacterial membrane protein insertase SecYEG-YidC

Max Busch[1,*], Cristian Rosales Hernandez [2, *], Michael Kamel [1, #, *], Yulia Schaumkessel [1, ¥], Eli van der Sluis [2, §], Otto Berninghausen [2], Thomas Becker [2], Roland Beckmann [2, ✉], Alexej Kedrov [1, 3, ✉]

[1] Synthetic Membrane Systems, Institute of Biochemistry, Heinrich Heine University Düsseldorf, Germany

[2] Gene Center Munich, Ludwig Maximilian University Munich, Germany

[3] Interfaculty Center for Membrane Research, Heinrich Heine University Düsseldorf, Germany

Current affiliations:

[#] Dept. Biology, Osnabrück University, Germany

[¥] Institute of Biochemistry and Molecular Biology I, Heinrich Heine University Düsseldorf, Germany

[§] Kavli Institute of Nanosciences, Technical University of Delft, the Netherlands

[*] These authors contributed equally to this work

✉ Corresponding authors

Abstract

The universally conserved Sec translocon and the YidC/Oxa1-type insertases mediate biogenesis of α-helical membrane proteins, but the molecular basis of their cooperation has remained disputed over decades. A recent discovery of a multi-subunit insertase in eukaryotes has raised the question about the architecture of the putative bacterial ortholog SecYEG-YidC and its functional mechanism. Here, we combine cryogenic electron microscopy with cell-free protein synthesis in nanodiscs to visualize biogenesis of the polytopic membrane protein NuoK, the subunit K of NADH-quinone oxidoreductase, that requires both SecYEG and YidC for insertion. We demonstrate that YidC is recruited to the back of the translocon at the late stage of the substrate insertion, in resemblance to the eukaryotic system, and in vivo experiments indicate that the complex assembly is vital for the cells. The nascent chain does not utilize the lateral gate of SecYEG, but enters the lipid membrane at the SecYE-YidC interface, with YidC being the primary insertase. SecYEG-YidC complex promotes folding of the nascent helices at the interface prior their insertion, so the examined cellular pathway follows the fundamental thermodynamic principles of membrane protein folding. Our data provide the first detailed insight on the elusive insertase machinery in the physiologically relevant environment, highlight the importance of the nascent chain for its assembly, and prove the evolutionary conservation of the gate-independent insertion route.

Introduction

Integral membrane proteins (IMPs) constitute up to 30 % of a genome content for each organism and determine the functionality of the cellular membranes[1]. Biogenesis of α-helical IMPs at the cytoplasmic membrane of bacteria and the endoplasmic reticulum (ER) in eukaryotes typically occurs in a co-translational manner: The nascent transmembrane helices (TMHs) emerging from the ribosome are identified by the signal recognition particle (SRP) and delivered in the form of ribosome-nascent chain complexes (RNCs) to the universally conserved Sec translocon that facilitates their insertion in the lipid bilayer [2,3]. The translocon of *Escherichia coli* (*E. coli*) is a heterotrimeric complex composed of SecY, SecE and SecG proteins forming a central pore. N- and C-terminal halves of the central subunit SecY may open laterally in a clam-shell manner, thus opening a lipid-exposed gap between the opposing TMHs 2b and 7. This gap, commonly referred as the “lateral gate”, has been seen as the exclusive site for the nascent TMHs to partition into the lipid phase, both for bacterial SecYEG and the homologous Sec61αβγ in eukaryotes. This model was supported by extensive biochemical studies [4,5] and more recently by structural insights, commonly focusing on signal peptides of secretory proteins and single-span IMPs exposing their N-terminal end into the cytoplasm [6-8]. However, polytopic IMPs have their TMHs in alternating topology, suggesting higher complexity within the biogenesis pathways [9,10].

Folding of numerous bacterial IMPs relies on a cooperative action between SecYEG and the essential and ubiquitous membrane protein insertase YidC [11]. YidC may distort the cytoplasmic lipid leaflet and facilitate folding and membrane insertion of the nascent IMPs [12,13]. Extensive intermolecular cross-linking studies suggested that YidC is located in front of the lateral gate of SecYEG where it could access the nascent TMHs emerging from the lateral gate [5,14], and the model was further supported by low-resolution cryo-EM analysis [15]. Remarkably, the AlphaFold model predicts a position of YidC of *E. coli* and other Gram-negative bacteria at the back of SecYEG, and diverse architectures are predicted for the homologs from Gram-positive bacteria, such as

Bacillus subtilis (Suppl. Figure 1). In all the cases, the model confidence is low, suggesting that other factors, such as the bound ribosome and/or the nascent chain may be required for the complex assembly.

Recent biochemical and high-resolution cryo-EM studies of the eukaryotic insertase machinery revealed that the YidC homolog TMCO1 was located at the back of the Sec61 complex, being a component of the large multi-subunit complex BOS-GEL-PAT [16,17]. Furthermore, insertion of a nascent polytopic IMP occurred not via the lateral gate, but between Sec61 and the BOS complex, thus challenging the long-standing paradigm in the field. The accumulated controversy between the insights from the eukaryotic homologs and the conventional model of the SecYEG-YidC complex has raised the question about the mechanism of IMP folding in bacteria and the architecture and dynamics of the insertion machinery. To tackle this issue, here we employed nanodisc-reconstituted SecYEG-YidC and cell-free protein synthesis (CFPS) to analyze the biogenesis of the polytopic IMP NuoK, the subunit K of the NADH-quinone oxidoreductase that requires both insertases for proper membrane integration and folding in *E. coli* [18]. By means of cryo-EM, we describe the detailed architecture of the RNC-bound SecYEG-YidC complex, where YidC is recruited to the back of SecYEG in the substrate-dependent manner thus indeed resembling the architecture of the eukaryotic insertase machinery. The nascent chain emerging from the SecY-bound ribosome is routed away from the lateral gate to the crossed SecY TMH 10 and SecE TMH 3 and then handed over to YidC to be inserted into the lipid bilayer via the dedicated groove. Overall, the data elucidate the functional architecture of the bacterial cotranslational IMP insertion machinery and reveal the evolutionary conservation of the “back-of-Sec" insertion route, thereby offering a new perspective on membrane protein biogenesis in bacteria.

Results

SecYEG and YidC-dependent insertion in nanodiscs

We set out to elucidate the architecture of the SecYEG-YidC complex by means of cryo-electron microscopy (cryo-EM) using lipid-based nanodiscs as a well-defined system for studying IMP insertion [6,19,20] (Figure 1A). To co-reconstitute both SecYEG and YidC in the relevant topology and to avoid their stochastic distribution among the nanodiscs, YidC and SecE were genetically fused via a cleavable linker (Suppl. Figure 2A), while the non-essential catioinic C-terminal end of YidC was removed to exclude spontaneous electrostatic interactions with the ribosome observed *in vitro* [19,21]. The fusion protein was expressed and isolated together with SecY and SecG subunits (Figure 1B). The complex was reconstituted into a sufficiently large nanodisc formed by MSP2N2 scaffold protein (expected diameter up to 17 nm) and treated with the HRV-3C protease to release YidC from SecE [22] (Figure 1C and Suppl. Figure 2B-D). As a result, each nanodisc contained one copy of both SecYEG and YidC, presumably in the correct orientation, while the dimensions of the surrounding lipid bilayer would allow for lateral diffusion of the proteins, assembly of the functional complex and insertion of the nascent TMHs. Supporting this notion, the molecular weight distribution of the nanodisc-embedded SecYEG-YidC complexes centered at 290 kDa (Figure 1D), and the difference from the calculated total protein weight of ~225 kDa likely reflected the presence of the co-reconstituted lipids.

We focused on biogenesis of *E. coli* NuoK (Suppl. Figure 3A), as both *in vivo* and *in vitro* experiments confirmed its dependence on SecYEG and YidC [18,23]. To ensure sequential incorporation of nascent NuoK TMHs into the membrane and reduce the risk of erroneous insertase:substrate complexes, we supplied the SecYEG-YidC nanodiscs to the CFPS reaction based on *E. coli* S30 extract. With that, the translating ribosomes were targeted to SecYEG-YidC, presumably by the endogenous SRP and the SRP receptor FtsY, to initiate nascent IMP insertion. To form a defined co-translational insertion intermediate, the gene construct coded for the first three TMHs of NuoK (NuoK^{86}) followed by a 31 aa long linker, incl. HA-tag and a C-terminal ribosome stalling sequence SecM* [24,25] (Suppl. Figure 3B and C). With this construct, after CFPS, the SecM* peptide and the linker/tag would occupy the exit tunnel of the stalled ribosome, while all three NuoK TMHs would be exposed and accessible for the insertase complex (Figure 1A).

Visualiation of the active RNC:insertase complex

The purified NuoK^{86}-RNC:SecYEG-YidC complexes were stabilized by rapid glutaraldehyde crosslinking before vitrification and cryo-EM analysis. After 3D classification followed by focused classification and refinement of the tunnel exit region, we obtained a structure of this complex with an average resolution of 2.4 Å (Suppl. Figure 4). The reconstruction showed a ribosome programmed with tRNAs in the A- and P-sites, a continuous nascent chain density ranging from the CCA-end of P-site tRNA to the tunnel exit and a clear density below the exit tunnel accounting for SecYEG and YidC embedded into a nanodisc of approx. 15 nm in diameter (Figure 1E and F). Focused refinement of the nanodisc density resulted in a local resolution of 3 to 5 Å for SecYEG and 4 to 7 Å for most parts of YidC (Suppl. Figure 4B and C) and allowed unambiguous docking of the modelled SecYEG-YidC complex with only minor adjustments (Suppl. Figure 5). The structure shows the RNC-bound SecYEG in the center of the nanodisc, while YidC is positioned at the back side of SecYEG near SecE TMH 3, with the large P1 domain of YidC exposed to the “periplasmic” side of the nanodisc. An additional density accounting for TMH 2 and TMH 3 of NuoK appears between SecYEG and YidC,

building extensive interactions with both proteins within the membrane and at the interface (Figure 1E). At lower contour levels, an additional transmembrane density appeared near NuoK TMH 2, which likely represents NuoK TMH 1 (Suppl. Figure 5F), though YidC TMH 1 could not be excluded.

As YidC was initially fused with SecE TMH 1 near the lateral gate, it must have relocated substantially within the lipid-filled nanodisc after the linker cleavage. To test whether the localization of YidC was not artificially induced by the chemical crosslinking, we carried out cryo-EM analysis of the non-crosslinked NuoK^{86}-RNC:SecYEG-YidC sample (Suppl. Figure 6). Though somewhat weaker and less well resolved, YidC density was clearly present in the same position at the back of SecYEG, implying that the crosslinked sample reflects the endogenous architecture of the substrate-bound SecYEG-YidC machinery.

Architecture of the SecYEG-YidC insertase

The clear density for all TMHs of SecY, SecE and SecG present in the local refined map offered the by now best-resolved view on the translocon structure in the physiologically relevant lipid environment (Figure 2A; Suppl. Figure 5). As observed previously for SecYEG-only assemblies, SecY is anchored via its cytoplasmic loops 6/7 and 8/9 to the tunnel exit of the large ribosomal subunit 50S (Figure 2B and C). Interestingly, SecY loop 6/7 is somewhat remodeled compared to the previously reported RNC-FtsQ:SecYEG assembly [6]. To avoid a clash with the NuoK nascent chain, the loop shifts away from the exit tunnel towards H6/7 and uL23 (Suppl. Figure 7A). Here, Tyr-258 is tightly accommodated in a pocket formed by bases A91, U92 and G93 of H7, while Arg-256 stacks on A63 of H6, and Arg-255 interacts with the loop of uL23 (His-70/Gly-71/Gln-72) (Figure 2B). In fact, this distortion in loop 6/7 leads to a global re-positioning of SecYEG with respect to the ribosome, with the 50S tilting by ~10° towards the N-terminal half of SecYEG (Suppl. Figure 7B). Contacts between the loop 8/9 with 50S are established via Lys-348 with the flipped-out base C490 of 23S rRNA helix H24 and a stacking interaction between SecY Arg-357 with A1392 (H53) that may be stabilized by a salt bridge with Glu-18 of the proximate ribosomal protein uL23 (Figure 2C). Further interactions are via the backbone of SecY Lys-347 with U1318 (H50) and Arg-340, Lys-364 with the flipped-out base A1535 of H59. The C-terminal cytoplasmic extension of SecY TMH 10 is bowed at Met-425/Ser-426 toward the SecY loop 8/9, so another contact is formed with the flipped-out base C490 of rRNA H24 (likely via Ala-436) (Figure 2D). As this SecY:ribosome contact was absent in earlier SecYEG:ribosome structures [6,7,9], we speculated that its formation is associated with the specific nascent chain routed for insertion. Notably, the lateral gate formed by TMHs 2b and 7 is in a tightly closed conformation indicating that it is not employed by the NuoK nascent chain, and the central pore is sealed by the "plug" domain, TMH 2a of SecY (Suppl. Figure 7B and C).

The conformation of YidC within the RNC-bound complex clearly deviates from the structure of isolated YidC obtained via X-ray crystallography [26] (Suppl. Figure 5E). The paddle domain, i.e. the hairpin built of helices CH1 and CH2 at the cytoplasmic interface, shifts by 1 nm to the periphery of the complex. There, it appears in close vicinity to the ribosomal protein uL24, but does not form a direct contact (Figure 2A). With that, it allows the passage of the nascent chain (Figure 2E), as described below. Furthermore, YidC TMH 6 is located at the periphery of the nanodisc, and its C-terminal extension is not required for the ribosome:SecYEG-YidC assembly, though this polypeptide is essential for binding to an RNC in absence of SecYEG *in vitro* [19], and it is also important for the homologous ribosome:Oxa1 interactions in mitochondria [27]. The large density at the periplasmic side matched the structure of the YidC P1 domain known from crystallography studies [26,28]. With respect to SecYEG, YidC is oriented with its TMH 3 facing TMH 3 of SecE, and the tip of the paddle domain faces the N-terminal end of SecY, in agreement with the efficient crosslinks observed in *E. coli* membranes between the residue 399 of YidC and SecY [14]. At the periplasmic side, the essential amphipathic helix EH1 and SecY TMH 5 (Figure 2E) form the only pronounced direct contact between these two proteins. In agreement with the AlphaFold2 model and our map in this region, this contact is mediated by SecY Arg-211 and YidC Asp-329 forming a salt bridge (discussed below). Apart from the salt bridge, SecYEG and YidC are held together mainly by the NuoK^{86} nascent chain, i.e. TMH 3 at the cytoplasmic interface and TMH 2 within the membrane region, suggesting that YidC is likely recruited to the RNC-bound SecYEG in a substrate-dependent manner.

Routing and insertion of the nascent NuoK

The nascent chain could be unambiguously traced from the CCA-end of P-site tRNA to the tunnel exit site within the ribosome (Figures 1F and 3A). All side chains of the SecM* stalling element ($_{111}$WWWPRIRGPP$_{120}$) were resolved with Pro-120 attached to the CCA-end of the A-site tRNA, while the preceding modified motif $_{117}$RGP$_{119}$ ($_{163}$RAG$_{165}$ in SecM) was coupled to the CCA-end of the P-site tRNA (Figure 3A). The conformation of the stalling peptide, tRNAs and critical bases in the peptidyl-transferase center were highly similar to the structure of SecM-stalled *E. coli* ribosome [29], indicating the same trapping mechanism as described for the RAG/P motif and its modified versions. Further, residues Arg-115 and Arg-117 interacted with PSU2504 and A2062, and the aromatic rings of the tryptophan motif ($_{112}$WWW$_{114}$) with bases A2059 and U2586 in the 23S rRNA (Figure 3B). Less resolved density accounting for the linker and the HA-tag continues from the constriction by uL4 and uL22 towards the tunnel exit (Figure 1F).

At the mouth of the tunnel exit, the NuoK nascent chain acquires a helical fold accounting for three C-terminal turns of its TMH 3 docked between the SecY loops 6/7, 8/9 and C-terminal TMH 10 extension and the tip of uL24 (Figure 3C). This newly formed arrangement of SecY creates a hydrophobic pocket in the otherwise hydrophilic environment of the tunnel exit and the solvent, that allows early folding and accommodation of a short hydrophobic TMH (Figure 3D). Strikingly, the emerging nascent chain is not routed into the central channel or the lateral gate of SecY. Once approaching the cytoplasmic funnel of SecY, NuoK TMH 3 density is kinked by approx. 120° towards the back side of the SecYEG, and so it acquires an interfacial topology laying over TMH 3 of SecE and TMH 10 of SecY (Figure 3E). These crossed TMHs form a sawhorse-like crevice at the membrane interface oriented toward YidC. In its turn, the paddle domain of YidC covers the nascent NuoK helix from the cytoplasmic side and completes a triangular window for the hydrophobic substrate. This observation is also consistent with previously proposed interactions of the evolutionary conserved paddle with substrates [13]. Like the SRP M-domain or Get3/TRC40, the YidC paddle contains a “methionine bristle” for client binding and indeed, our structure now shows how the membrane-facing Met-408 and Met-409 in CH2 interact with NuoK TMH 3 prior its insertion (Figure 3E). Notably, replacing these methionine residues with lysines, but not alanine, rendered a cold-sensitive phenotype in *E. coli* upon depletion of the wild-type YidC (Figure 3F, Suppl. Figure 8). We speculate that the effect originates from defects in IMP biogenesis, and that the apolar environment provided by the paddle domain is an important factor for the nascent IMP folding at the lipid membrane interface prior the insertion, in agreement with the physical considerations [30].

The close association of YidC with SecYEG ensures immediate hand-over of the nascent chain to the YidC insertase, and the position of NuoK TMH 2 in our structure elucidates the insertion route (Figure 4A). The membrane-inserted NuoK TMH 2 is separated from TMH 3 by the kink, likely the short flexible loop, in a close contact with YidC TMH 3 at the level of Cys-423 of the insertase. NuoK TMH 2 occupies the wide groove between YidC TMHs 3 and 5, i.e. the functional insertase site described in previous studies[20,31]. As the groove narrows towards the periplasm, the N-terminal end of NuoK TMH 2 is expelled from the helical bundle of YidC, being fully exposed to the lipid environment. Notably, the visualized NuoK does not reach YidC TMH 2 and does not interact with Arg-366 located deep in the YidC groove (Figure 4A). This residue was shown to be important for YidC functionality, as it destabilizes the lipid membrane and potentially interacts with charged/polar insertion substrates [12,32]. Although not observed in our structure, we cannot exclude that this interaction is formed at other stages of NuoK insertion, or it may be specific for particular nascent IMPs.

Determinants for the SecYEG-YidC assembly

Since the structure of SecYEG-YidC is organized around the NuoK insertion intermediate, we next asked how specific elements of the nascent chain may affect the complexes’ assembly. As YidC involvement in NuoK biogenesis was previously linked to Glu-36 and Glu-72 residues within TMH 2 and TMH 3 of the nascent IMP, respectively [18], we substituted both glutamates with lysines and performed cryo-EM analysis of the NuoK^{86mut}-RNC:SecYEG-YidC complexes (Figure 4B; Suppl. Figure 9). Despite the charge inversion and without the glutaraldehyde stabilization, the YidC density was observed in the same position at the back of SecYEG as upon wild-type NuoK insertion. However, when the length of the nascent chain was reduced to 48 amino acids, so only TMH 1 and TMH 2 of NuoK were exposed from the ribosome (NuoK^{48}-RNC), only the RNC-bound SecYEG was resolved within the nanodisc. The signal for YidC was lost, so the insertase remained mobile within the surrounding membrane, and the nanodisc manifested the characteristic tilt by ~10° (Figure 4C; Suppl. Fig. 10). Thus, we concluded that the SecYEG-YidC complex is dynamically assembled to mediate insertion of a nascent IMPs emerging via the “back of Sec” route, and the YidC recruitment may not depend on specific charge within the substrate, but may rely on other determinants, such as the length and the folding status of the emerging chain.

Interestingly, the sequence-based prediction of the SecYEG-YidC assembly by AlphaFold showed a very similar arrangement of YidC versus SecYEG as observed in our insertion intermediate in nanodiscs, despite the absence of the ribosome and the nascent chain. Here, the salt bridge between SecY Arg-211 (TMH 5) and YidC Asp-329 (end of EH1) may be a key interaction between the two proteins (Figure 4D and Suppl. Fig. 11A). This salt bridge is found among multiple distant Gram-negative species, but it is absent in Gram-positives, such as *B. subtilis*, which correlated with differences in the AlphaFold predictions (Suppl. Figure 1). This further suggested that the charged interaction is not essential but may serve to stabilize the dynamic complex. Thus, we speculated that disrupting this interaction may be detrimental for IMP biogenesis and affect the cell viability. Accordingly, we tested the importance of the salt bridge by replacing YidC Asp-329 with either alanine or arginine. Both YidC variants could be recombinantly expressed and purified, suggesting that the mutated proteins were correctly folded (Suppl. Figure 11B and C). When testing the functional complementation of these YidC variants in *E.coli* FTL10 strain, we observed that the plasmid-encoded YidC^{D329A} fully supported the cell growth under depletion of the wild-type YidC (Figure 4E; Suppl. Figure 11D). Strikingly, introducing a positive

charge was detrimental for the cellular viability, similar to the known effect of a minor deletion within the proximate EH1 [20]. Thus, we concluded that the salt bridge indeed plays a stabilizing role upon the assembly of the dynamic SecYEG-YidC complex, while the electrostatic repulsion rendered via the YidC^{D329R} mutation prohibits YidC docking, causing defects in IMP folding, either at the insertion or intramembrane assembly stage.

Comparison of bacterial and eukaryotic insertases

Guiding of the NuoK nascent chain to the back side of SecYEG for YidC-mediated insertion manifests an obvious resemblance to the route of multi-spanning IMP biogenesis in the ER of eukaryotes [17]. Here, the ribosome-bound Sec61 translocon serves as a core for the substrate-dependent assembly of a super-complex including BOS (back-of-Sec61) and GEL (GET- and EMC-like) complexes. We thus compared the arrangement of the *E. coli* SecYEG-YidC insertase with the human Sec61-BOS-GEL-PAT machinery trapped with a nascent chain of rhodopsin[17] (Figure 5). The bacterial insertase forms a compact assembly, where YidC is tightly packed against SecYEG and the inserted part of NuoK, providing a simple solution for transferring the nascent chain from the ribosome to the insertion-competent groove of YidC. In contrast, the loosely packed eukaryotic super-complex extends far from the Sec61 core and the resolved nascent chain, and it forms a lipid-filled “peninsula”. Remarkably, the YidC homolog TMCO1 (a constituent of the GEL complex) is found more than 3 nm away from Sec61 with its paddle domain oriented towards the rest of GEL. The interaction with the core translocon is mediated via the BOS subunit TMEM147 that largely overlaps with the position of YidC in the bacterial machinery. Interestingly, also the route of the nascent chain differs between two structures: While the membrane-inserted opsin TMH 3 and the RNC-bound linker were found in the orthogonal geometry to the membrane without interactions with the BOS-GEL subunits (Figure 5B), we report here that the TMHs of NuoK required distinct YidC domains for insertion and folding at the interface. The involvement of YidC at different stages of NuoK biogenesis correlates with its dual role as insertase and as chaperone for nascent IMPs.

The structure of the eukaryotic machinery suggests that blockade of the lateral gate by the CCDC47 protein of the PAT complex is a possible determinant to guide a nascent chain to the alternative exit [17]. Here, the extended helices of CCDC47 at the cytoplasmic interface reach the N-terminal half of Sec61α, the SecY homolog, and restrict its mobility required for opening the gate (Figure 5B), so the nascent IMP is forced to follow the Sec61-GEL-BOS insertion route. Within the bacterial insertase, the C-terminal end of SecY at the RNC:SecYEG interface may play a similar role, though using a different mechanism. In presence of the NuoK TMH 3 emerging from the tunnel mouth, this SecY extension forms a contact with the rRNA base C490 and hinders the access to the lateral gate. Thus, it may favor the alternative route for the growing nascent chain to the back of SecYEG. Strikingly, shortening of the SecY C-terminus and deletion of Lys-433 and Lys-434 involved in RNA binding appeared lethal for the cells[33], highlighting the importance of this poorly studied element of SecYEG.

Discussion

Variations in TMH topologies and lengths, appearance of partially unfolded regions and charged residues within the membrane core, point to the complexity in IMP biogenesis across the kingdoms of life. Membrane insertion of hydrophobic TMHs via the lateral gate of the SecYEG/Sec61$\alpha\beta\gamma$ translocon has been described as a general route towards IMP folding and assembly. The universal paradigm has been lately challenged by the discovery of the eukaryotic insertase machinery, the novel multi-subunit complex BOS-GEL-PAT, that facilitates IMP insertion at the back of the ribosome-bound Sec61 away from the lateral gate. The discovery of a novel insertion route in eukaryotes has immediately raised a question about the evolutionary conservation of the mechanism. Here, we describe how the bacterial machinery composed of SecYEG and YidC insertases mediates the biogenesis of the multi-spanning IMP NuoK. Cryo-EM of the lipid-embedded complex has revealed the substrate-dependent recruitment of YidC to the ribosome-bound SecYEG and visualized the route taken by the nascent chain from the ribosome to the lipid membrane. Differently to the lactose permease folding intermediates assessed in a recent study [10], the NuoK nascent chain does not pass the lateral gate, but it folds within the hydrophobic pocket formed by SecY loops and leaves via the sawhorse-shaped crevice at the back, where it is encountered and inserted by YidC. Here, the conserved and partially hydrophobic paddle domain of YidC assists in handover of the pre-folded nascent chain from SecYEG to YidC. The SecYEG-YidC assembly is induced by the nascent chain of sufficient length, and it is stabilized by the salt bridge between SecY and YidC at the periplasmic interface. Although somewhat different in details and complexity level, the overall architecture of the active SecYEG-YidC complex manifests a conceptual similarity to the eukaryotic Sec61-BOS-GEL-PAT insertase machinery. Thus, our results not only validate the existence of the alternative insertion pathway in bacteria and reveal the path of the nascent chain but also strongly suggest an evolutionary link between the distant systems.

Multiple studies have described interactions of nascent IMPs with the lateral gate of SecYEG and suggested that YidC is located in proximity to access the emerging nascent TMHs [5,14,10]. Our findings do not exclude such architecture, though no relevant classes were observed in the acquired cryo-EM data

sets in this study. It remains to be explained what factors determine the route of nascent IMPs, i.e. the triage between the lateral gate and the "back-of-Sec" insertion, and also what features of the nascent chain serve as signals for YidC recruitment. For NuoK biogenesis, the necessity of YidC was previously linked to the anionic residues within the substrate, e.g. via interactions with YidC Arg-366 [18]. However, we show that YidC is docked at the back of SecYEG also when the charges within the NuoK TMHs are inverted, while no contact with Arg-366 is observed. Notably, the involvement of YidC at the late stage of NuoK biogenesis correlates with the recent findings that both YidC and the evolutionary related EMC complex facilitate insertion of the C-terminal TMHs of polytopic IMPs [34,35]. This final insertion event would occur post-translationally when the translation termination happens soon after the synthesis of the last TMH. NuoK would be an example of such IMP: In absence of the stalling introduced here for the purpose of the structural analysis, the nascent polypeptide chain would be released from the peptidyl transferase center, and the membrane insertion of NuoK TMH 3 complex would occur post-translationally. As both YidC and EMC ensure folding of post-translationally delivered clients [16,36,37], the C-terminal TMHs of polytopic IMPs may utilize the same recognition and insertion routes, while the SecYEG-YidC assembly would facilitate the efficient substrate delivery.

The topology acquired by NuoK in the nanodisc is another puzzling outcome of our study. The model based on the resolved TMHs suggests that the N-terminus of the insertion intermediate is oriented into the cytoplasm, thus being inverted in comparison to NuoK within the assembled complex I (Suppl. Figure 3A). One explanation may be that the non-native topology originated from the experimental *in vitro* set-up, as the employed nanodiscs lack the physiological trans-membrane electrostatic potential, an important factor for the IMP orientation. However, if the electrostatic potential was a decisive factor, we would expect to find NuoK in two different topologies in cryo-EM reconstructions, possibly utilizing different insertion routes, while only one class was experimentally observed. Thus, an alternative scenario seems plausible: Several IMPs, such as four-TMH EmrE, were shown to undergo the complete topology inversion in the last step of their folding, and a single charged residue here can be decisive for the final orientation [38,39,40]. As NuoK does not contain long loops between the TMHs but contains a stretch of three arginine residues within its C-terminal end, it is tempting to speculate that the protein acquires its correct topology in the cellular membrane once the complete polypeptide chain is released from the ribosome. The flip-flop of the inserted TMHs may be promoted by the proximate YidC, as the insertase would distort the lipid bilayer and so destabilize the NuoK intermediate that eventually results in the new topology.

Our study delivers a new, direct view on the organization of the bacterial SecYEG-YidC insertase machinery, and the revealed architecture is calling for detailed investigation on its dynamics, as well as re-evaluation of the earlier data. One current limitation of the employed nanodisc-based system is the pre-defined stoichiometry of SecYEG and YidC within the complex, while YidC is substantially more abundant in the bacterial membrane [41]. Thus, even larger insertase complexes may be envisioned for the cellular membrane, where multiple YidC copies could manage biogenesis of polytopic IMPs, resembling the Sec-BOS-GEL machinery of the eukaryotes (Figure 5). Experiments on the native bacterial membranes, such as those by cryogenic electron tomography [42], as well as reconstituted systems will be of the utmost importance to explore such modes of interaction and deliver detailed and possibly direct insights on the key process of membrane protein biogenesis.

Acknowledgements

The authors would like to thank Hanna Kratzat, Susanne Rieder and Charlotte Ungewickel for the support with the cryo-EM collection and initial data analysis, Joana Musial for assistance with cloning and biochemical characterization of the SecYEG-YidC construct, Laura Czech and Gert Bange for the support upon the development of CFPS procedure, and Arnold J.M. Driessen and Florian Altegoer for discussions. The work was funded by German Research Foundation (DFG, grants Ke1879/3 and Collaborative Research Center 1208, project A10 to A.K.), European Research Council (Advanced Grant "CryoTranslation" to R.B.)

Contributions

MB: biochemical sample preparation and analysis, *in vivo* experiments; CRH: cryo-EM data processing, model building and visualization; MK: biochemical sample preparation and analysis; YS: CFPS development; EvdS: design and characterization of the SecYEG-YidC insertase; OB: cryo-EM data collection and curation; TB: cryo-EM data analysis and visualization; RB and AK: project conceptualization and supervision, funding acquisition. The initial manuscript was prepared by MB, CRH and AK, and edited/commented by all the co-authors.

Additional information

Supplementary information is available for this paper.

Correspondence and requests should be addressed to Roland Beckmann (Beckmann@genzentrum.lmu.de) and Alexej Kedrov (Kedrov@hhu.de).

Figures

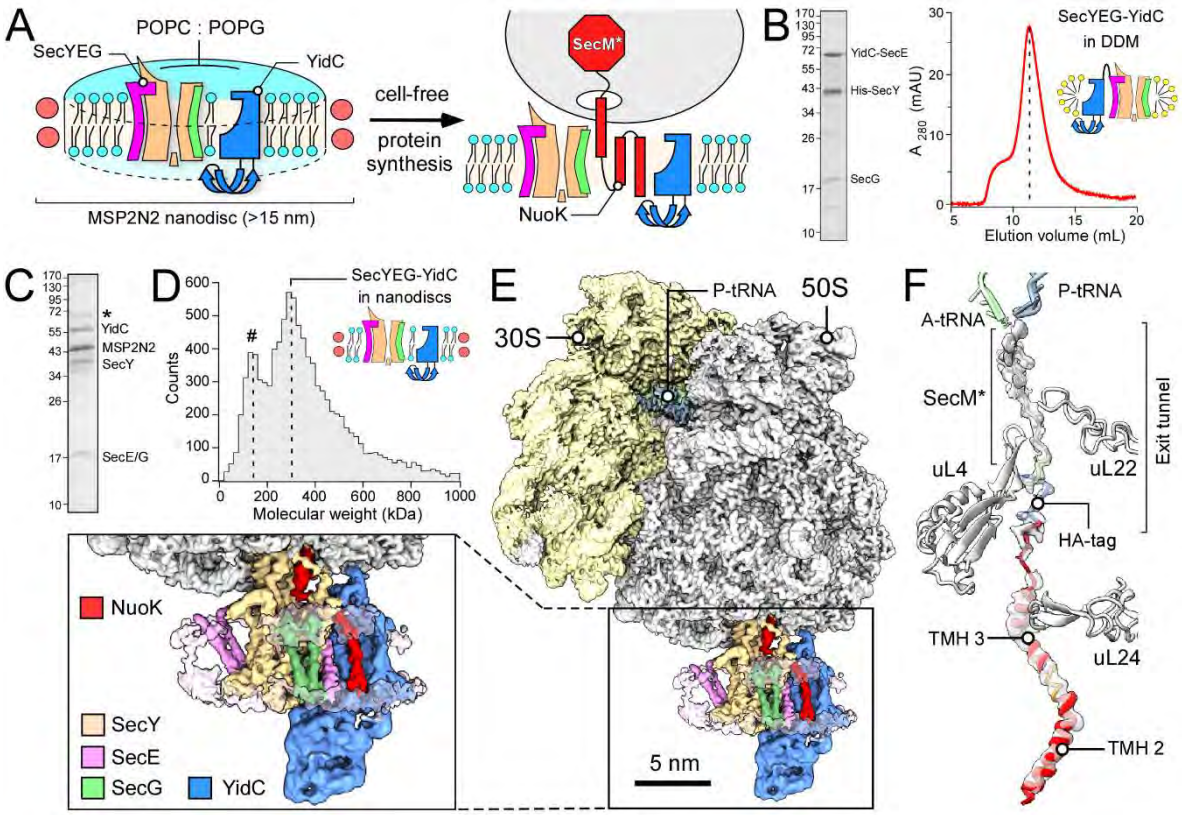

Figure 1. Substrate-induced assembly of the RNC:SecYEG-YidC complex.

(A) Scheme of the assay to study NuoK biogenesis. SecYEG and YidC are co-reconstituted into MSP2N2-based nanodisc in presence of POPC.POPG lipids and introduced into CFPS reaction. Synthesis of the substrate NuoK is interrupted by SecM* stalling sequence, so the stable insertion intermediate is formed for structural analysis.

(B) SDS-PAGE of the affinity-purified SecYEG-YidC fusion insertase complex, with the subunits indicated. Right: Size exclusion chromatography profile of the isolated insertase complex.

(C) SDS-PAGE of the nanodisc-reconstituted SecYEG-YidC complex after the protease treatment. The minor band at 70 kDa (*) corresponds to the residual YidC-SecE fusion protein.

(D) Mass photometry recording of isolated SecYEG-YidC nanodiscs manifests the main peak at 290 kDa. The minor peak at 130 kDa (#) corresponds to the nanodiscs loaded only with lipids.

(E) Cryo-EM reconstruction of NuoK^{86}-RNC:SecYEG-YidC assembly. Shown is a composite map consisting of isolated densities of the ribosome subunits 30S and 50S, P-site peptidyl-tRNA ("P-tRNA") and focused refined SecYEG-YidC, with subunits and the substrate NuoK indicated.

(F) Isolated density of the nascent chain NuoK (in transparent) with fitted model; tRNAs and the ribosomal proteins lining the exit tunnel are indicated.

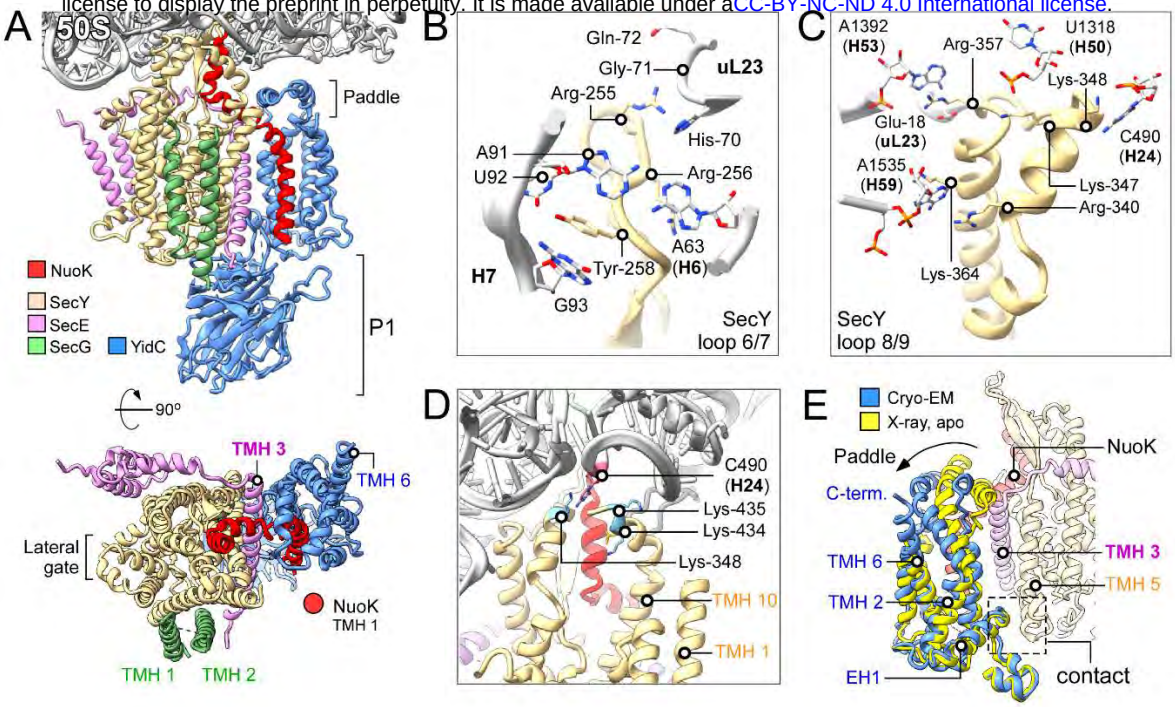

Figure 2. Architecture of SecYEG-YidC insertase and its interactions with the RNC.

(**A**) Molecular model of the SecYEG-YidC insertase complex bound to NuoK^{86}-RNC. The putative position of the membrane-inserted NuoK TMH 1 is indicated as a red circle in the cytoplasmic view (bottom).

(**B, C**) The network of ribosome:SecYEG interactions mediated by SecY cytoplasmic loops 6/7 and 8/9.

(**D**) The C-terminal extansion of SecY contacts the flipped-out base C490 of 23S rRNA. This interaction is likely stabilized by Lys-435 close to the H24 rRNA backbone and Lys-348 stacking on the C490 from the opposite site.

(**E**) Overlay of X-ray and cryo-EM structures of YidC highlight the conformational change within the insertase. The movement of the YidC paddle domain upon entry of the NuoK nascent chain is indicated with an arrow. The closest SecY-YidC contact at the periplasmic side is indicated with the dashed box.

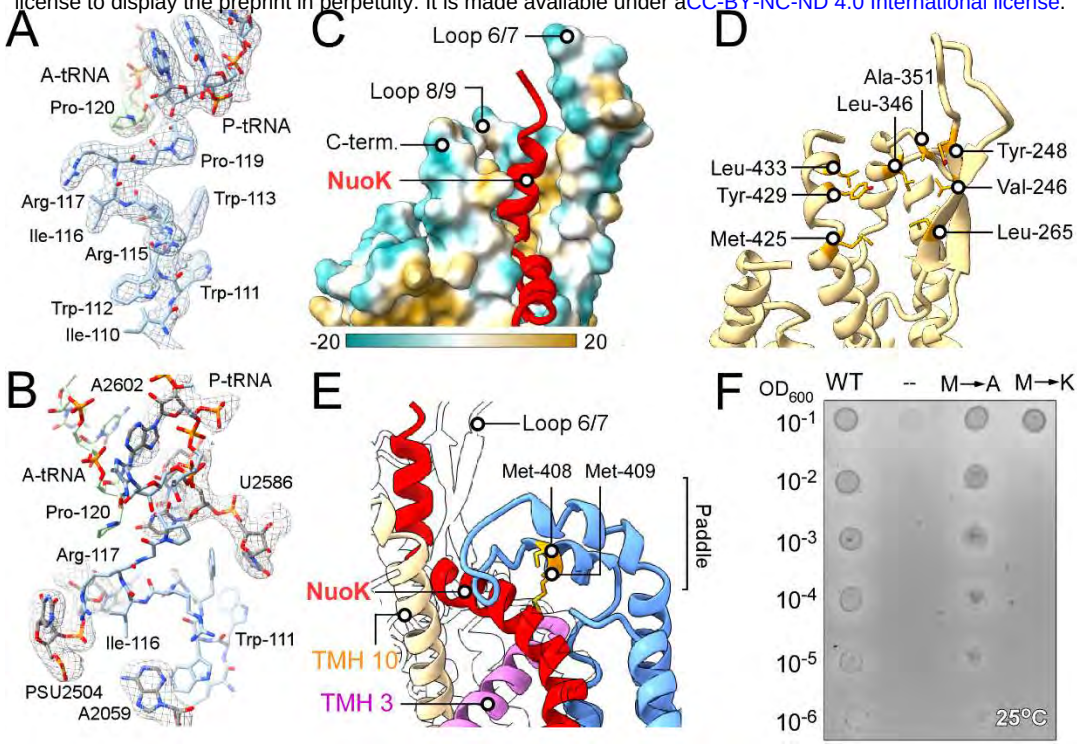

Figure 3. **SecYEG and YidC jointly mediate delivery of the nascent chain to the membrane**.

(**A**) Isolated cryo-EM density (transparent mesh) of the SecM* stalling sequence in the ribosomal exit tunnel with the fitted model. The residues of SecM* peptide are indicated.

(**B**) Model for the SecM*:ribosome interactions. Hallmark bases of the PTC and involved in SecM* binding are shown as in SecM-stalled ribosome (PDB 8QOA), fitted into the respective density in (transparent mesh).

(**C**) Surface representation of the molecular model for SecY cytoplasmic loops 6/7, 8/9 and the C-terminal tail colored according to the molecular lipophilicity potential (scale bar shown). The emerging NuoK nascent chain (red ribbon) acquires α-helical fold within the hydrophobic pocket.

(**D**) The indicated apolar residues form the hydrophobic pocket for the nascent chain folding.

(**E**) Zoom view on the molecular model for the SecYEG-YidC complex highlighting the route of the folded NuoK TMH 3 (red ribbon) from SecYE to YidC. The methionine residues within the YidC paddle domain are indicated.

(**F**) Complementation test using wild-type YidC ("WT"), mutants $YidC^{M408A,\ M409A}$ ("M→A") and $YidC^{M408K,\ M409K}$ ("M→K") and empty vector ("--") in presence of glucose at 25°C.

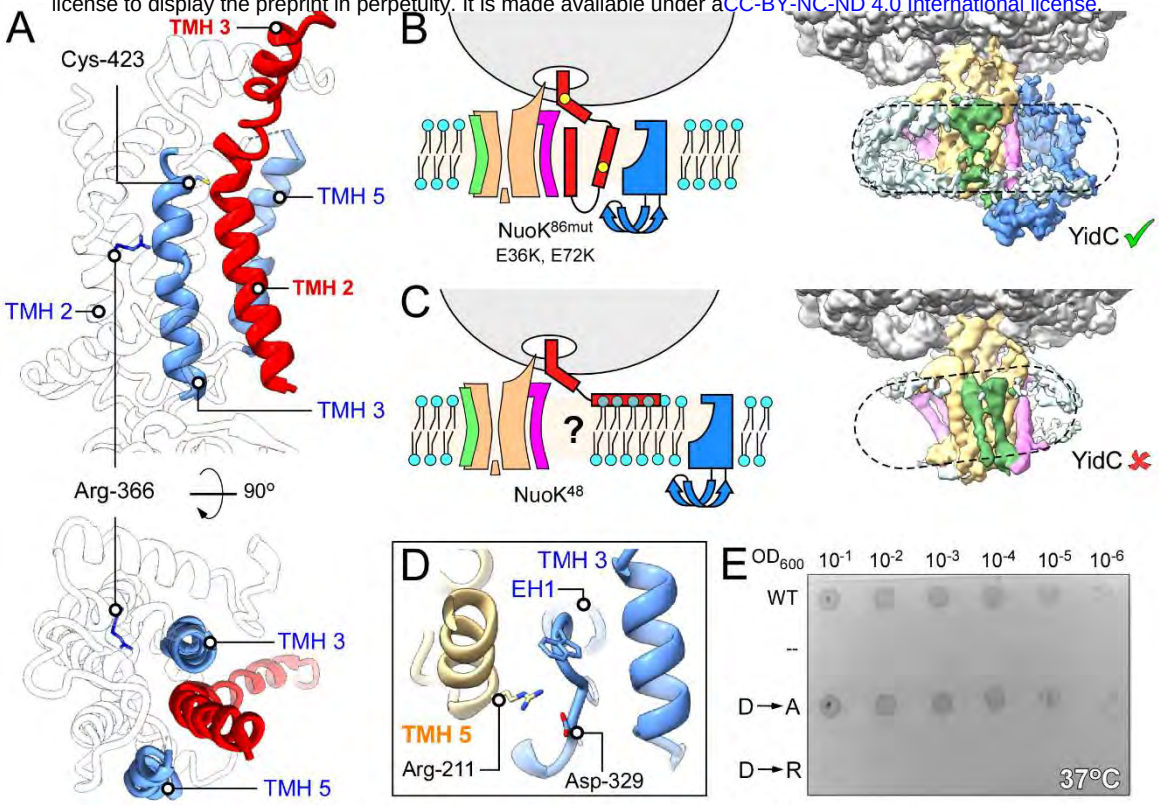

Figure 4. YidC-mediated insertion of the nascent membrane protein.

(**A**) View on the molecular model of YidC focusing on the insertion groove between TMHs 3 and 5 with the accommodated NuoK TMH2. Arg-366 within YidC TMH 2 is not involved in interactions with the nascent chain. The rest of YidC is shown in transparent.

(**B**) Insertion model of the NuoK^{86} intermediate bearing E36K, E72K mutations (yellow circles) and the corresponding cryo-EM map (right). The outline of the nanodisc is shown with the dashed line.

(**C**) The insertion model of the NuoK^{48} intermediate and the corresponding cryo-EM map (right). The outline of the nanodisc is shown with the dashed line. The color code in (**B**) and (**C**) is as in Figure 1E.

(**D**) View on the molecular model of the SecY-YidC complex focusing on the closest contact site at the periplasmic interface. A putative salt bridge is formed between SecY Arg-211 and YidC Asp-329.

(**E**) Complementation test using wild-type YidC ("WT"), mutants YidC^{D329A} ("D→A") and YidC^{D329R} ("D→R") and empty vector ("--") in presence of glucose at 37°C.

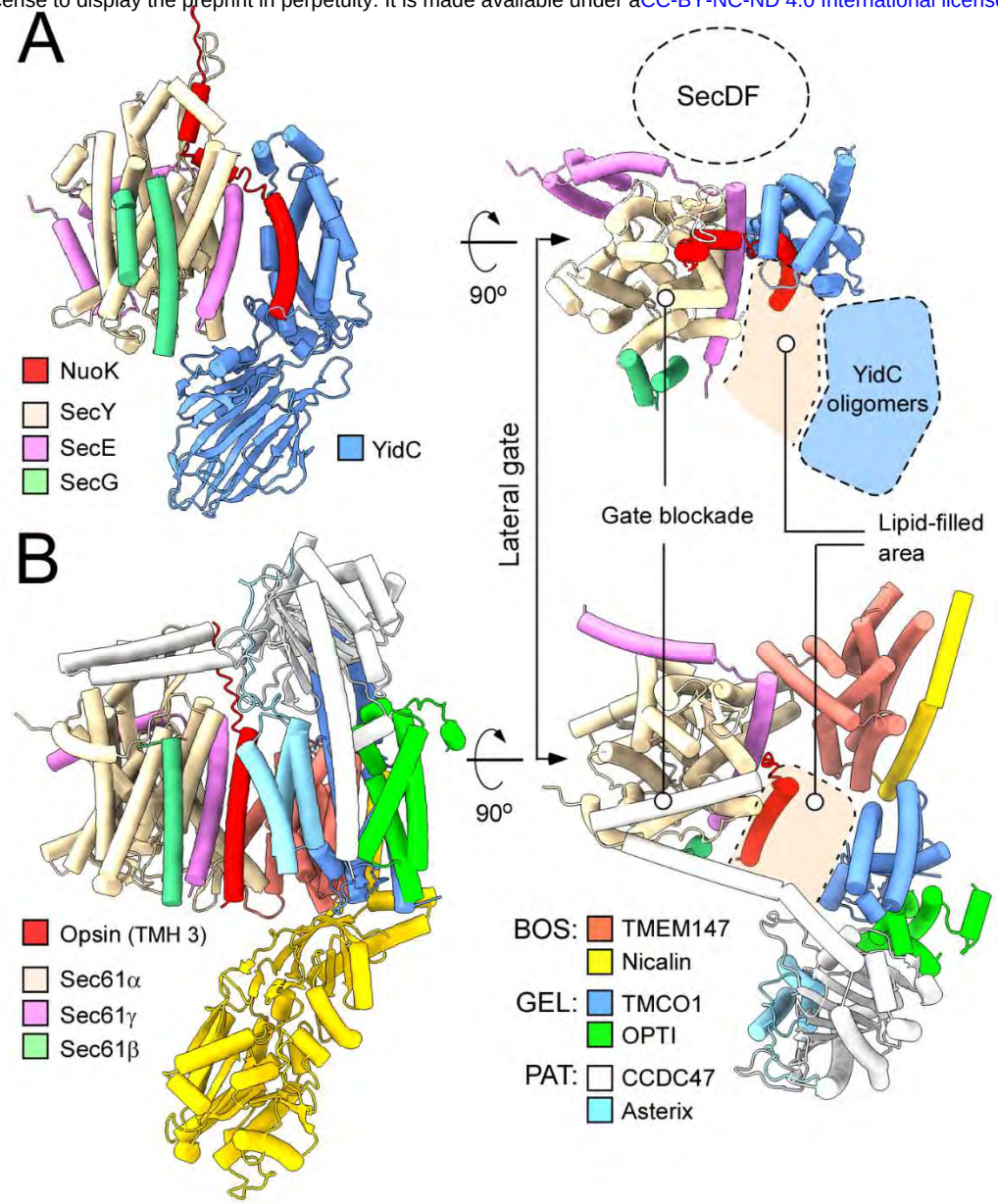

Figure 5. Comparison of the bacterial and eukaryotic IMP insertases.

(A) Side and cytoplasmic views on *E. coli* SecYEG-YidC assembly and the route of the NuoK nascent chain. The putative position of the accessory SecDF complex is indicated, based on a high-confidence AlphaFold3 prediction. A hypothetical position of YidC oligomers is shown in blue.

(B) Side and cytoplasmic views on the eukaryotic insertase with the subcomplexes and individual subunits indicated. The partially resolved route of the opsin nascent chain is indicated. Similar structural and functional features of the bacterial and eukaryotic machineries are shown.

Methods

Molecular cloning

The primers were synthesized by Eurofins Genomics and Merck/Sigma-Aldrich. Sequencing was conducted by Eurofins Genomics and Microsynth AG. The used restriction enzymes and cloning kits, incl. the Phusion High-Fidelity DNA polymerase and the Gibson assembly kits were purchased by New England Biolabs.

The genes encoding for SecYEG-YidC complex were cloned in a pBAD-TOPO (Thermo Fisher Scientific) derived vector (plasmid ID pEM472). The gene encoded for *E. coli* YidC (1-535) was fused to the N-terminus of SecE with a sequence coding for HRV-3C protease cleavage site inserted in between; the construct also contained a N-terminally decahistidine-tagged SecY and unmodified SecG. The construct was created using conventional restriction-ligation cloning techniques. For the IMP substrate, the DNA encoding 86 amino acids of *E. coli* NuoK and the C-terminal extension containing the HA-tag for immunodetection, a linker and the SecM* stalling sequence (FSTPVWIWWWPRIRGPP) (Ref. [29]) was synthesized by GenScript Biotech (Netherlands), and the construct was cloned cloned into pRSET vector (Thermo Fisher Scientific) via SalI/HindIII restriction sites. The gene encoding NuoK^{48} nascent chain was prepared by removing the fragment corresponding to NuoK residues 49-86 via PCR, followed by blunt-end ligation.To study the effects of point mutations within YidC in complementation assays, the YidC mutants were cloned into pTrc99A-based plasmid pKAD107 [19]. For the expression of the YidC mutants, the corresponding genes were cloned into pBAD-based plasmid [20].

Expression of the membrane insertases

E.coli C41(DE3)Δ*ompF*-Δ*acrAB* strain [43] transformed with the pEM472 plasmid was grown in LB medium supplemented with 100 µg/mL ampicillin at 37°C while shaking at 180 rpm. Upon reaching OD_{600} 0.6, expression of the fused SecYEG-YidC complex was induced by adding 0.5 % arabinose (w/v) and carried out for 2.5 h. Cells were harvested and lysed using a microfluidizer (M-110P, Microfluidics Corp.), and the debris was removed by centrifugation at 18000 g for 10 min (SS34 rotor, Sorvall/Thermo Scientific). Membranes were isolated by ultracentrifugation at 42000 rpm for 1 h (45Ti rotor, Beckman Coulter). Membranes were then suspended in the solubilization buffer (50 mM Hepes-KOH pH 7.4, 150 mM KOAc, 5 % glycerol, 200 µM tris(2-carboxyethyl)phosphine (TCEP), 1 % n-dodecyl β-maltoside (DDM; Glycon Biochemicals GmbH) for 1 h at 4°C. The solubilized material was centrifuged at 21,380 x *g* for 10 min at a tabletop centrifuge (Hermle Z 216 MK, Hermle Labortechnik GmbH) and the supernatant was incubated with Ni^{2+}-NTA agarose resin (Macherey-Nagel GmbH & Co. KG) for 1 h at 4°C. The resin was washed with 50 mM Hepes-KOH pH 7.4, 500 mM KOAc, 5 % glycerol (v/v), 200 µM TCEP, 0.05 % DDM, 10 mM imidazole and the target protein was eluted with 50 mM Hepes-KOH pH 7.4, 150 mM KOAc, 5 % glycerol, 200 µM TCEP, 0.05 % DDM, 300 mM imidazole. The elution fractions were concentrated and subjected to size exclusion chromatography (SEC) in 50 mM Hepes-KOH pH 7.4, 150 mM KOAc, 5 % glycerol, 0.05% DDM using Superdex 200 Increase GL 10/300 column with ÄKTA Pure set-up (Cytiva). The protein concentration was determined spectrophotometrically (NeoDot, NeoBiotech) using calculated extinction coefficient of 134,000 M^{-1} cm^{-1}. Expression of YidC mutants, $YidC^{D329A}$ and $YidC^{D329R}$, was achieved using pBAD-based plasmid and the proteins were solubilized and purified in DDM according to the previously published protocol [20].

Assembly of SecYEG-YidC nanodiscs

To prepare the nanodiscs, liposomes were first formed using 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC; 70 mol %) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1-rac-glycerol) (POPG; 30 mol %) (Avanti Polar Lipids, Inc). The lipids were mixed from chloroform stocks to achieve the desired ratio, the solvent was evaporated and the liposomes were prepared, as described [6]. The detergent-purified SecYEG-YidC complexes were reconstituted in MSP2N2-based nanodiscs in a protein:MSP:lipid molar ratio of 1:4:400 following the published protocol [6]. After forming the nanodiscs, the linker connecting YidC and SecE was cleaved by HRV-3C protease for 1 h at 4°C and SEC was performed in in 50 mM HEPES-KOH pH 7.4, 150 mM KOAc, 5 % glycerol using Superose 6 Increase GL 10/300 column with ÄKTA Pure set-up. SEC fractions containing SecYEG-YidC nanodiscs were concentrated to ~10 µM using Amicon Ultra-4, Ultracel 30 K centrifugal filters (Merck/Millipore). Mass photometry measurements were performed using Two^{MP} instrument (Refeyn Ltd.) calibrated in the range from 66 to 1,048 kDa with NativeMark Unstained Protein Standard (Invitrogen/Thermo Fisher Scientific).

Cell-free protein synthesis

E. coli S30 extract for CFPS was prepared based on previously published protocols [44]. Briefly, *E.coli* BL21(DE3) cells were transformed with TargoTron pAR1219 plasmid (Sigma-Aldrich) encoding for T7 RNA polymerase. 2 L of 2x YPTG media was inoculated with 100 mL overnight cultures, the cells were grown to OD_{600} 0.5, and the T7 RNA polymerase expression was induced with 1 mM IPTG. The cells were further grown to reach OD_{600} 1.0 and then they were harvested at 7500 rpm for 15 min (FiberLite F8-6x1000y rotor, Piramoon Technologies Inc.). The cell pellet was washed three times with 10 mM Tris-acetate pH 8, 60 mM KOAc, 14 mM $Mg(OAc)_2$, and 1 mM PMSF, and the pellet was resuspended in the same buffer at the ratio of 1 mL per 1 g pellet. Subsequently, cells were lysed by sonication (10 times, 15 s on, 30 s off, 50 % power, 5 pulsed cycles) (Sonopuls GM2200, Bandelin). The lysate was cleared via two-steps

centrifugation, 12,000 x *g* for 15 min and 30,000 x *g* for 30 min (S120 AT6 rotor, Sorvall /Thermo Fisher). The supernatant was aliquoted and stored at -75°C.

The CFPS reaction was composed of 40 % S30 extract, the master mix (10 mM ammonium acetate, 130 mM KOAc, 33 mM sodium pyruvate, 1.5 mM spermidine, 1 mM putrescine, 4 mM sodium oxalate, 1.2 mM ATP, 0.85 mM of GTP, CTP and UTP, 34 µg/ml folinic acid, 170.6 µg/mL of *E.coli* tRNA MRE 600 (Roche Diagnostics GmbH), 0.33 mM NAD^+, 0.26 mM coenzyme A and 2 mM of each amino acid), and $Mg(OAc)_2$. The optimum $Mg(OAc)_2$ concentration was identified for each new batch of the S30 extract upon screened within the range of 0 to 12 mM and using synthesis of the yellow fluorescent protein as a read-out. For NuoK synthesis, at least 7 ng/µL plasmid DNA encoding the nascent chain was added to the reaction, as well as at least 100 nM of the nanodisc-reconstituted SecYEG-YidC. CFPS reactions were performed at 37°C for 1 h while shaking at 450 rpm. The synthesis/stalling was evaluated via Western-blotting using monoclonal antibodies against the HA-tag (sc-7392, Santa Cruz Biotechnology).

Isolation of ribosomes from CFPS reactions

To isolate the ribosomes, 10-40 % linear sucrose gradients were formed in SW40 tubes (Beckman Coulter) using the Gradient Station (BioComp Instruments). CFPS reactions (100 µL) were loaded on top of the gradients and the samples were centrifuged for 16 h at 16,500 rpm (SW40 Ti rotor; Beckman Coulter). The gradients were fractionated using the Gradient Station while monitoring the absorbance at 280 nm (A_{280}). The peaks with the ribosomal fractions occurring in the sucrose concentration range of 20-25 % were pooled together and incubated with Ni^{2+}-NTA agarose resin (Macherey-Nagel GmbH & Co. KG) for 1 h at 4°C. The resin was washed with 50 mM Hepes-KOH pH 7.4, 500 mM KOAc, 25mM $Mg(OAc)_2$, 10 mM imidazole. The complexes were eluted with 50 mM Hepes-KOH pH 7.4, 150 mM KOAc, 25 mM $Mg(OAc)_2$, 300 mM imidazole and concentrated using an Amicon Ultra-4 Ultracel 30 K centrifugal filters (Merck/Millipore) while exchanging the buffer to 50 mM Hepes-KOH pH 7.4, 150 mM KOAc, 25 mM $Mg(OAc)_2$. The presence of the nascent chain with ribosomes was confirmed by a western blot against the HA-tag. The concentration of ribosomes was then estimated by measuring the absorbance at 260 nm (A_{260}).

In order to improve the resolution of the cryo-EM data on $NuoK^{86}$-RNC:SecYEG-YidC, glutaraldehyde was added to the purified samples. The sample was diluted 5-fold to prevent inter-particle crosslinking, and it was then incubated with 0.1 % glutaraldehyde (v/v) for 15 min on ice and quenched by 100 mM Tris-HCl pH 7.5. Afterwards the sample was concentrated using a Amicon Ultra-4, Ultracel 30 K centrifugal filters, flash-frozen and stored at -75°C before grid preparation.

Cryo-EM sample preparation and data collection

Isolated NuoK-RNC:SecYEG-YidC samples were supplemented with (1H, 1H, 2H, 2H-perfluoro-octyl)-β-D-maltopyranoside (FOM, Anatrace) to a final concentration of 0.03% to favor random orientation of the particles and plunge frozen. For each grid, 3.5 µL of the sample was applied onto glow-discharged Quantifoil Cu 300 mesh R3/3 grids with an additional 2 nm layer of carbon. After a waiting time of 45 s, the grids were blotted for 3 s and plunge frozen in liquid ethane at 4°C and 100% humidity using a Vitrobot Mark IV (Thermo Fisher Scientific). Data collection for RNC-NuoK:SecYEG-YidC samples was performed at 300 keV using a Titan Krios microscope equipped with a Falcon 4i direct electron detector and a SelectrisX imaging filter (all Thermo Fisher Scientific) at a pixel size of 0.727 Â. Dose-fractioned movies were collected in a defocus range from -0.5 to 3.0 µm and with a total dose of 40 e^- per $Å^2$, fractionated in 40 frames to obtain a total dose of 1 e^- per $Å^2$ per frame. Gain correction, movie alignment and summation of movie frames was performed using MotionCor2 [45]. Further processing, including CTF estimation, was carried out in cryoSPARC v4.4 [46].

Data processing

For the crosslinked $NuoK^{86}$-RNC:SecYEG-YidC complex, 27,660 micrographs were selected. Blob Picker was used to pick 1,420,623 particles which were sorted by 2D classification, yielding a subset of 636,995 particles. An ab-initio job with 2 classes was run to further clean the particle set. A consensus refinement of 582,663 particles resulted in a map of 70S with clear extra density below the tunnel exit. A soft mask covering this region (accounting for nanodisc, SecY/YidC and eventually the NuoK nascent chain) was used to sort the particles into 6 classes, using a 3D Classification job. Two of the classes, representing a total of 221,144 particles, displayed a strong SecYEG and YidC density. The class with best resolved nascent chain density containing 113,368 particles (19.5 %) was selected and refined to a final resolution of 2.44 Å. Local refinement was performed on the tunnel exit region yielding a map with a final resolution of 3.76 Å. This map was used to build the atomic model for the NuoK:SecYEG-YidC assembly (PDB: 9RBF; EMDB:XXX; EMDB:XXX (local refinement); EMDB:XXXX (composite map)). Data processing for this dataset is summarized in Suppl. Figure 4.

For the non-crosslinked $NuoK^{86}$-RNC:SecYEG-YidC assembly, 10,234 micrographs were selected and manually curated. A total of 693,102 particles was picked with Blob Picker and sorted by 2D classification to generate templates for the Template Picker job. From 2,300,222 template-picked particles, 406,978 were selected after 2D classification. An ab-initio job (three classes) allowed further cleaning of the data set. A consensus refinement with a set of 217,418 particles resulted in a map of 70S ribosome with clear extra density below the tunnel exit. Focused sorting into five classes in the tunnel exit region yielded a class of

33,912 particles showing density for the extramembrane P1 domain of YidC. These particles were used to train a TOPAZ picking model, which resulted in 360,610 particles. The particles were further curated and focused sorted into a final class of 70,670 particles. This class was refined to a final resolution of 2.73 Å (EMD-53587). A local refinement in the SecYEG-YidC region yielded a map with a final resolution of 3.19 Å (EMD-53589). Data processing for this dataset is summarized in Suppl. Figure 6.

For the NuoK^{86mut}-RNC:SecYEG-YidC complex assembled upon NuoK$^{E36K,\ E72K}$ synthesis Blob picker was used to pick an initial set of 586,721 particles from 9,099 micrographs. The particles were curated and after 2D classification, a subset of 287,399 particles was used to generate an ab-initio reconstruction (two classes) of 70S ribosomes with extra density for the insertase complex. The same set was further used to generate templates for template picking, which resulted in 2,057,403 particles, from which a final set of 56,980 particles was obtained after extensive 2D and 3D classification. The particles were fed into a TOPAZ training job, obtaining 433,249 particles after picking and extraction. These particles were used to generate a consensus refinement and focused sorted into eight classes. The class with the best resolved insertase region, comprising 48,952 particles, was further refined to generate a reconstruction with final resolution of 2.87 Å (EMD-53584). A local refinement with a mask in the translocon region generated a map of 3.20 Å (EMD-53585). Data processing for this dataset is summarized in Suppl. Figure 9.

For the NuoK^{48}-RNC:SecYEG-YidC complex assembled upon synthesis of the early NuoK intermediate, an initial set of 741,311 particles was picked using a Blob picker job based on 21,440 micrographs. The particles were cleaned by 2D classification jobs to yield a final 511,016 particle set. An ab-initio reconstruction (two classes) job was used to generate an initial map of the ribosome-bound SecYEG in the nanodisc. Successive, focused 3D-classifications from a consensus refinement were used to obtain a final class of 66,486 particles. These particles were refined to a final resolution of 2.62 Å (EMD-53560). A local refinement with a mask in the translocon region resulted in a reconstruction of 2.90 Å (EMD-53568). Data processing for this dataset is summarized in Suppl. Figure 10.

For visualization of all cryo-EM maps, consensus maps were generated consisting of isolated density for the 70S ribosome from the global refinements and isolated density for the nanodisc-embedded insertase from the local refinements.

Model building and refinement

A molecular model was built for the crosslinked NuoK^{86}-RNC:SecYEG-YidC complex. For the 70S ribosome a previously released model based of high-resolution cryo-EM maps of a 70S ribosome and a SecM-stalled RNC were used as templates (PDB IDs 7K00, 8QOA) 29,47. The model for the tRNA-Gly (in SecM) was changed to tRNA-Pro and the mRNA model was adusted from GCU-GGC-CCU (Ala-Gly-Pro in SecM) to GGU-CCU-CCG (Gly-Pro-Pro in SecM*). A model for the SecYEG-YidC assembly was generated based on the rigid body fitting the AlphaFold3 prediction of this complex (see also Suppl. Figure 4) that was fitted into the locally refined cryo-EM density with only minor adjustments (see also Suppl. Figure 5). For the nascent chain, the modified SecM* sequence could be modelled de novo based on well-resolved density for this region (Figue 3A). Less-resolved density was present to fit the backbone and a few bulky side chains for major parts of glycine-serine linker and the HA-tag (Figure 1F). The NuoK TMH 2 and 3 were identified based on the rod-like shape of extra density present at the mouth of the exit tunnel and between SecY and YidC (Figure 1F). While at the given resolution we cannot be sure about the exact register, we positioned residues 73-82 of NuoK THM 3 into the corresponding density supporting the predeicted α-helical conformation. 70S ribosome and SecYEG-YidC atomic models were processed independently in multiple rounds of manual real-space refinement in Coot v0.9.8.95 [49]. The models were later merged and further refined in Phenix v1.20.1-4487 [50]. The Molprobity tool was used for model validation. Visualization was done in UCSF ChimeraX v1.9 [51].

Complementation assay

The complementation assay was prepared as described before [19]. Briefly, a single colony of FTL10 cells [52] transformed with either YidC-encoding plasmid or empty pTrc99A vector was grown in LB medium with 0.2% arabinose, 25 µg/mL kanamycin and 100 µg/mL ampicillin for 16 h at 37°C at 180 rpm. The overnight cultures were diluted to an OD_{600} of 0.05 and grown until the early logarithmic phase before diluting them all to an OD_{600} of 0.1 and doing a subsequent serial dilution. 5 µL of each diluted culture were transfered on the plates (LB agar supplemented with 25 µg/mL kanamycin, 100 µg/mL ampicillin and either 0.2% arabinose or 0.2% glucose). The plates were incubated for 16 h at 37°C or 24 h at 25°C.

Data availability

The cryo-EM structural data have been deposited in the Electron Microscopy Data Base (EMDB) and the Protein Data Bank (PDB) repositories under the accession numbers EMD-53892, EMD-53893, EMD-53894 and PDB-9RBF (NuoK^{86}-RNC:SecYEG-YidC (crosslinked); EMD-53587 and EMD-53589 (NuoK^{86}-RNC:SecYEG-YidC); EMD-53584 and EMD-53585 (NuoK^{86mut}-RNC:SecYEG-YidC); EMD-53560 and EMD-53568 (NuoK^{48}-RNC:SecYEG-YidC).

References

1 von Heijne, G. The membrane protein universe: what's out there and why bother? *J Intern Med* **261**, 543-557 (2007). https://doi.org/10.1111/j.1365-2796.2007.01792.x

2 Itskanov, S. & Park, E. Mechanism of Protein Translocation by the Sec61 Translocon Complex. *Cold Spring Harb Perspect Biol* **15** (2023). https://doi.org/10.1101/cshperspect.a041250

3 Cymer, F., von Heijne, G. & White, S. H. Mechanisms of integral membrane protein insertion and folding. *J Mol Biol* **427**, 999-1022 (2015). https://doi.org/10.1016/j.jmb.2014.09.014

4 du Plessis, D. J., Berrelkamp, G., Nouwen, N. & Driessen, A. J. The lateral gate of SecYEG opens during protein translocation. *J Biol Chem* **284**, 15805-15814 (2009). https://doi.org/10.1074/jbc.M901855200

5 Sachelaru, I. *et al.* YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein. *J Biol Chem* **288**, 16295-16307 (2013). https://doi.org/10.1074/jbc.M112.446583

6 Kater, L. *et al.* Partially inserted nascent chain unzips the lateral gate of the Sec translocon. *EMBO Rep* **20**, e48191 (2019). https://doi.org/10.15252/embr.201948191

7 Jomaa, A., Boehringer, D., Leibundgut, M. & Ban, N. Structures of the E. coli translating ribosome with SRP and its receptor and with the translocon. *Nat Commun* **7**, 10471 (2016). https://doi.org/10.1038/ncomms10471

8 Voorhees, R. M. & Hegde, R. S. Structure of the Sec61 channel opened by a signal sequence. *Science* **351**, 88-91 (2016). https://doi.org/10.1126/science.aad4992

9 Bischoff, L., Wickles, S., Berninghausen, O., van der Sluis, E. O. & Beckmann, R. Visualization of a polytopic membrane protein during SecY-mediated membrane insertion. *Nat Commun* **5**, 4103 (2014). https://doi.org/10.1038/ncomms5103

10 Ou, X. *et al.* SecY translocon chaperones protein folding during membrane protein insertion. *Cell* **188**, 1912-1924 e1913 (2025). https://doi.org/10.1016/j.cell.2025.01.037

11 Steinberg, R., Knupffer, L., Origi, A., Asti, R. & Koch, H. G. Co-translational protein targeting in bacteria. *FEMS Microbiol Lett* **365** (2018). https://doi.org/10.1093/femsle/fny095

12 Chen, Y. *et al.* YidC Insertase of Escherichia coli: Water Accessibility and Membrane Shaping. *Structure* **25**, 1403-1414 e1403 (2017). https://doi.org/10.1016/j.str.2017.07.008

13 McDowell, M. A., Heimes, M. & Sinning, I. Structural and molecular mechanisms for membrane protein biogenesis by the Oxa1 superfamily. *Nat Struct Mol Biol* **28**, 234-239 (2021). https://doi.org/10.1038/s41594-021-00567-9

14 Petriman, N. A. *et al.* The interaction network of the YidC insertase with the SecYEG translocon, SRP and the SRP receptor FtsY. *Sci Rep* **8**, 578 (2018). https://doi.org/10.1038/s41598-017-19019-w

15 Botte, M. *et al.* A central cavity within the holo-translocon suggests a mechanism for membrane protein insertion. *Sci Rep* **6**, 38399 (2016). https://doi.org/10.1038/srep38399

16 Page, K. R. *et al.* Role of a holo-insertase complex in the biogenesis of biophysically diverse ER membrane proteins. *Mol Cell* **84**, 3302-3319 e3311 (2024). https://doi.org/10.1016/j.molcel.2024.08.005

17 Smalinskaite, L., Kim, M. K., Lewis, A. J. O., Keenan, R. J. & Hegde, R. S. Mechanism of an intramembrane chaperone for multipass membrane proteins. *Nature* **611**, 161-166 (2022). https://doi.org/10.1038/s41586-022-05336-2

18 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). https://doi.org/10.1074/jbc.M109.051128

19 Kedrov, A. *et al.* Elucidating the native architecture of the YidC: ribosome complex. *J Mol Biol* **425**, 4112-4124 (2013). https://doi.org/10.1016/j.jmb.2013.07.042

20 Kedrov, A. *et al.* Structural Dynamics of the YidC:Ribosome Complex during Membrane Protein Biogenesis. *Cell Rep* **17**, 2943-2954 (2016). https://doi.org/10.1016/j.celrep.2016.11.059

21 Kohler, R. *et al.* YidC and Oxa1 form dimeric insertion pores on the translating ribosome. *Mol Cell* **34**, 344-353 (2009). https://doi.org/10.1016/j.molcel.2009.04.019

22 Ritchie, T. K. *et al.* Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods Enzymol* **464**, 211-231 (2009). https://doi.org/10.1016/S0076-6879(09)64011-8

23 Price, C. E. & Driessen, A. J. YidC is involved in the biogenesis of anaerobic respiratory complexes in the inner membrane of Escherichia coli. *J Biol Chem* **283**, 26921-26927 (2008). https://doi.org/10.1074/jbc.M804490200

24 Cymer, F., Hedman, R., Ismail, N. & von Heijne, G. Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. *J Biol Chem* **290**, 10208-10215 (2015). https://doi.org/10.1074/jbc.M115.641555

25 Kempf, N. *et al.* A Novel Method to Evaluate Ribosomal Performance in Cell-Free Protein Synthesis Systems. *Sci Rep* **7**, 46753 (2017). https://doi.org/10.1038/srep46753

26 Kumazaki, K. *et al.* Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. *Scientific Reports* **4** (2014). https://doi.org/10.1038/srep07299

27 Haque, M. E., Spremulli, L. L. & Fecko, C. J. Identification of protein-protein and protein-ribosome interacting regions of the C-terminal tail of human mitochondrial inner membrane protein Oxa1L. *J Biol Chem* **285**, 34991-34998 (2010). https://doi.org/10.1074/jbc.M110.163808

28 Ravaud, S., Stjepanovic, G., Wild, K. & Sinning, I. The crystal structure of the periplasmic domain of the Escherichia coli membrane protein insertase YidC contains a substrate binding cleft. *J Biol Chem* **283**, 9350-9358 (2008). https://doi.org/10.1074/jbc.M710493200

29 Gersteuer, F. *et al.* The SecM arrest peptide traps a pre-peptide bond formation state of the ribosome. *Nat Commun* **15**, 2431 (2024). https://doi.org/10.1038/s41467-024-46762-2

30 White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu Rev Biophys Biomol Struct* **28**, 319-365 (1999). https://doi.org/10.1146/annurev.biophys.28.1.319

31 Yuan, J., Phillips, G. J. & Dalbey, R. E. Isolation of cold-sensitive yidC mutants provides insights into the substrate profile of the YidC insertase and the importance of transmembrane 3 in YidC function. *J Bacteriol* **189**, 8961-8972 (2007). https://doi.org/10.1128/JB.01365-07

32 Kumazaki, K. *et al.* Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* **509**, 516-520 (2014). https://doi.org/10.1038/nature13167

33 Chiba, K., Mori, H. & Ito, K. Roles of the C-terminal end of SecY in protein translocation and viability of Escherichia coli. *J Bacteriol* **184**, 2243-2250 (2002). https://doi.org/10.1128/JB.184.8.2243-2250.2002

34 Wu, H., Smalinskaite, L. & Hegde, R. S. EMC rectifies the topology of multipass membrane proteins. *Nat Struct Mol Biol* **31**, 32-41 (2024). https://doi.org/10.1038/s41594-023-01120-6

35 Kalinin, I. A. *et al.* Features of membrane protein sequence direct post-translational insertion. *Nat Commun* **15**, 10198 (2024). https://doi.org/10.1038/s41467-024-54575-6

36 Anghel, S. A., McGilvray, P. T., Hegde, R. S. & Keenan, R. J. Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum. *Cell Rep* **21**, 3708-3716 (2017). https://doi.org/10.1016/j.celrep.2017.12.006

37 Hennon, S. W., Soman, R., Zhu, L. & Dalbey, R. E. YidC/Alb3/Oxa1 Family of Insertases. *J Biol Chem* **290**, 14866-14874 (2015). https://doi.org/10.1074/jbc.R115.638171

38 Seurig, M., Ek, M., von Heijne, G. & Fluman, N. Dynamic membrane topology in an unassembled membrane protein. *Nat Chem Biol* **15**, 945-948 (2019). https://doi.org/10.1038/s41589-019-0356-9

39 Seppala, S., Slusky, J. S., Lloris-Garcera, P., Rapp, M. & von Heijne, G. Control of membrane protein topology by a single C-terminal residue. *Science* **328**, 1698-1700 (2010). https://doi.org/10.1126/science.1188950

40 Woodall, N. B., Hadley, S., Yin, Y. & Bowie, J. U. Complete topology inversion can be part of normal membrane protein biogenesis. *Protein Sci* **26**, 824-833 (2017). https://doi.org/10.1002/pro.3131

41 Urbanus, M. L. *et al.* Targeting, insertion, and localization of Escherichia coli YidC. *J Biol Chem* **277**, 12718-12723 (2002). https://doi.org/10.1074/jbc.M200311200

42 Gemmer, M. *et al.* Visualization of translation and protein biogenesis at the ER membrane. *Nature* **614**, 160-167 (2023). https://doi.org/10.1038/s41586-022-05638-5

43 Kanonenberg, K. *et al.* Shaping the lipid composition of bacterial membranes for membrane protein production. *Microb Cell Fact* **18**, 131 (2019). https://doi.org/10.1186/s12934-019-1182-1

44 Zubay, G. In vitro synthesis of protein in microbial systems. *Annu Rev Genet* **7**, 267-287 (1973). https://doi.org/10.1146/annurev.ge.07.120173.001411

45 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017). https://doi.org/10.1038/nmeth.4193

46 Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017). https://doi.org/10.1038/nmeth.4169

47 Watson, Z. L. *et al.* Structure of the bacterial ribosome at 2 A resolution. *Elife* **9** (2020). https://doi.org/10.7554/eLife.60482

48 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024). https://doi.org/10.1038/s41586-024-07487w

49 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004). https://doi.org/10.1107/S0907444904019158

50 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221 (2010). https://doi.org/10.1107/S0907444909052925

51 Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018). https://doi.org/10.1002/pro.3235

52 Hatzixanthis, K., Palmer, T. & Sargent, F. A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol Microbiol* **49**, 1377-1390 (2003). https://doi.org/10.1046/j.1365-2958.2003.03642.x