

Supplementary Information

**The role of the N-terminal amphipathic helix in bacterial YidC:  
Insights from functional studies, the crystal structure and molecular  
dynamics simulations**

Karol J. Nass<sup>a,1</sup>, Ioana M. Ilie<sup>b,1,\*</sup>, Manfred J. Saller<sup>c,2</sup>, Arnold J.M. Driessen<sup>c</sup>, Amedeo Caflisch<sup>b</sup>, Richard A. Kammerer<sup>d</sup> and Xiaodan Li<sup>d,\*</sup>

- a. Photon Science Division, Paul Scherrer Institute, Forschungstrasse 111, 5232 Villigen PSI, Switzerland
- b. Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057, Zurich, Switzerland
- c. Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, Nijenborgh 7, 9727 AG Groningen, The Netherlands
- d. Laboratory of Biomolecular Research, Division of Biology and Chemistry, Paul Scherrer Institute, Forschungstrasse 111, 5232 Villigen PSI, Switzerland.

\*Corresponding authors: [i.m.ilie@uva.nl](mailto:i.m.ilie@uva.nl); [xiao.li@psi.ch](mailto:xiao.li@psi.ch)

1: contributed equally

2: current affiliation: Lava Therapeutics NV, Yalelaan 60, 3584CM Utrecht, The Netherlands

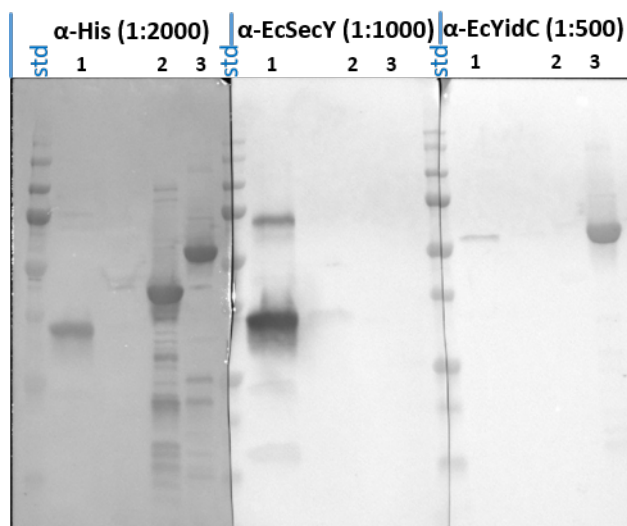


Figure S1: Interaction of EcSecY and EcYidC. Purified EcSecY, EcYidC and TmYidC all containing a hexa-his tag run on Bio-Rad anyKd gel and then blotted on nitrocellulose membrane. Three primary antibodies were used to detect his-tagged protein (anti penta-His. 1:2000), EcSecY (anti-SecY rabbit, SN810-2, 1:5000) and EcYidC (anti-EcYidC rabbit, self-made, 1:500). The secondary antibody for penta-His is goat-anti-mouse (1:4000) conjugated with Alkaline-Phosphatase and the secondary antibody for EcSecY and EcYidC is donkey anti-rabbit (1:4000) conjugated with alkaline phosphatase. The blot was developed 30 sec for colorimetric detection.

```

T.maritima : MVLRKVVAILLAILPIFLF : 19
E.coli      : MDSQRNLLVIALLFVSFMI : 19
S.enterica : MDSQRNLLVIALLFVSFMI : 19
Y.pestis   : MDSQRNLLLIALLFVSFMI : 19
L.pneumoph : MDIRRIVLYMALALIGLSL : 19
K.pneumoni : MDSQRNLLIIALLFVSFMI : 19
H.influenz : MDSRRSLLVLALIFISFLV : 19
N.meningit : MDFKRLTAFFAIALVIMIG : 19

```

Figure S2: The N-terminal amphipathic helix (N-AH) sequence alignment of different Gram-negative bacteria using ClustalX (1).

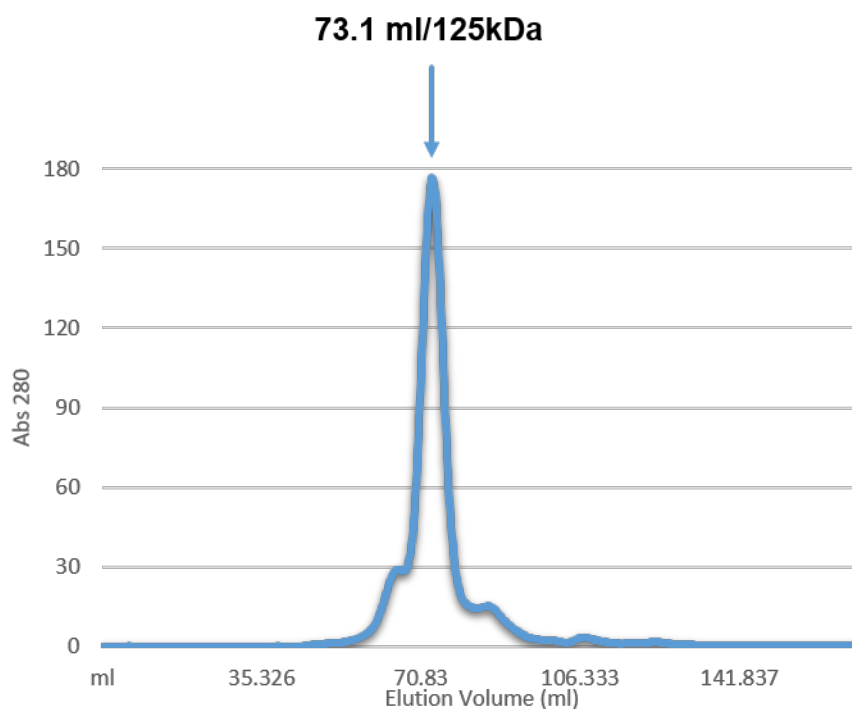


Figure S3: Gel filtration elution profile of purified TmYidC in the presence of 0.05%DDM. Purified TmYidC loaded on a Superdex 200 16 60 column using 50mM Hepes, pH 7.5, 150mM and 0.05% DDM buffer. The arrow pointed the elution volume of TmYidC and its molecular weight calculated based on the elution volume using the Bio-Rad standard running at the same condition.

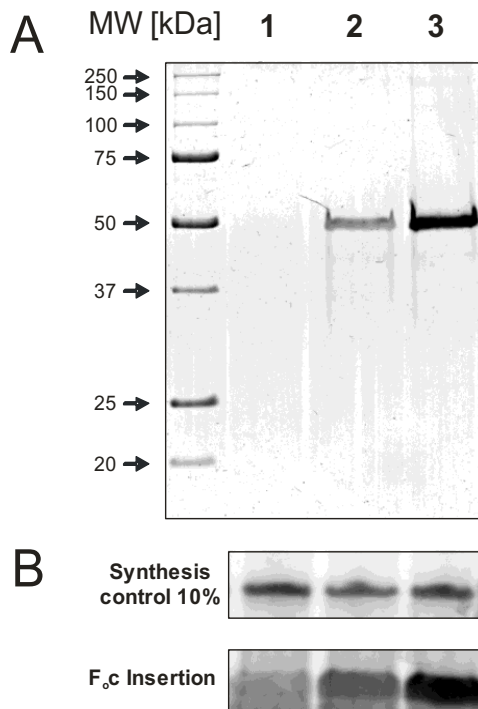


Figure S4: *Thermotoga maritima* YidC mediates the membrane insertion of *in vitro* synthesized *E. coli* F<sub>0</sub>C. TmYidC was overexpressed in *E. coli* SF100 and affinity purified as described. (A) The purified protein was reconstituted in *E. coli* total lipids at different protein to lipid ratios and analyzed by SDS

PAGE. Lane 1: no protein; lane 2: protein/lipid ratio 0.08 (wt/wt), lane 3: protein/lipid ratio 0.16 (wt/wt). Proteoliposomes were tested for the membrane insertion of *in vitro* synthesized <sup>35</sup>S labeled *E. coli* F<sub>0</sub>C. The total F<sub>0</sub>C synthesis was equal in the presence of different proteoliposomes (B, upper panel), while the amount of membrane inserted F<sub>0</sub>C (lower panel) increased with the TmYidC concentration in the reaction mixture.

### Strains and plasmids

The sequences of all primers used in this study are listed in Table S1. All strains and plasmids are summarized in Table S2. In order to express TmYidC and EcYidC/TmYidC hybrid genes in *E. coli* cells, genes were cloned into the *E. coli* expression vector pTrc99A (Table S2). Constructs were verified by DNA sequencing. For expression of TmYidC with a His<sub>6</sub> tag attached in frame to its 3' end, the gene was amplified from genomic DNA of *T. maritima* MSB8 (generous gift from S.V. Albers) (2) using the primer pair FoTmaNcoI/ReTmHisXbaI (Table S1). The PCR product was cleaved with NcoI and XbaI and ligated into the corresponding sites of the vector pTrc99A, yielding pTrcTmHis (Table S2).

The EcYidC and TmYidC hybrids were constructed by overlapping PCR. The first (5') 186, 963 and 1038 bps of *yidC* were amplified from genomic DNA of *E. coli* DH5α using the forward primer FoYidCXbaI and either ReYidC61Tma396, ReYidC321Tma225 or ReYidC346Tma198 as reverse primer (Table S1). Subsequently, DNA fragments consisting of the last (3') 1188, 675 and 594 bps of *tm1461* were amplified from genomic DNA of *T. maritima* MSB8 using either FoYidC61Tm396, FoYidC321Tm225 or FoYidC346Tm198 as forward and ReTmHisHindIII as reverse primer. The fragments 186 and 1188, 963 and 675, and 1038 and 594 were fused by overlapping PCR using the combined DNA fragments as template and the primer pair FoYidCXbaI/ReTmHisHindIII. The PCR products were cleaved with XbaI and HindIII and ligated into the corresponding sites of pTrc99A, yielding pTrce61-t396His, pTrce321-t225His and pTrce346-t198His.

Table S1. Primers used in this study

Primer	Sequence (5'→3')
FoTmNcoI	GCTGCCATGGTCTTGAGAAAAGTTGTAGC
ReTmHisXbaI	GCGTCTAGATCAGTGATGGTGATGGTGATGTGCCTTTTTCGGAAGACCAAGAAGTTCTCTC
FoYidCXbaI	GCTCTAGATGGATTCGCAACGC
ReTmHisHindIII	CACGAAGCTTTCAGTGATGGTGATGGTGATGTGCC
FoYidC61Tm396	GTGGCCAGGGGAAACTGATCTCGGGCATACTCAAAGACTTTTACACCCTT
ReYidC61Tm396	CGAGATCAGTTTCCCCTGG
FoYidC321Tm225	GACAAAATGGCAGCTGTTGCTCCAGGATTCAACAAGTGGTA
ReYidC321Tm225	TACCACTTGTTGAATCCTGGAGCAACAGCTGCCATTTTGTC
FoYidC346Tm198	TGTTCAAACCTGCTGAAATGGTTCGGATGGGCGATCATGCT
ReYidC346Tm198	AGCATGATCGCCATCCGAACCATTTTCAGCAGTTTGAACA

Table S2. Strains and plasmids used in this study

Plasmids or strain	Relevant properties	Reference
<b>Plasmids</b>		
pTrc99A	Expression vector for <i>E. coli</i> based on pKK233-2, carries a hybrid <i>trp/lac</i> promoter and the multiple cloning site of pUC18; Ap <sup>R</sup>	(3)
pTrcyidC	Like pTrc99A, contains <i>E. coli yidC</i>	(4)
pTrcTmaHis	Like pTrc99A, contains <i>T. maritima tm1461</i> with a His <sub>6</sub> tag at 3' end	This study
pTrce61-t396His	Like pTrc99A, contains the 3' 183 bps of <i>E. coli yidC</i> fused in frame to the 5' 1188 bps of <i>T. maritima tm1461</i> with a His <sub>6</sub> tag at 3' end	This study
<b><i>E. coli</i> strains</b>		
DH5α	<i>supE44 hsdR14 recA1 endA1 gyrA96 thi-1 relA1 ΔlacU169 (Φ80lacZ ΔM15)</i> ; K12 derivative	(5)
SF100	F <sup>-</sup> <i>ΔlacX74 galK thi rpsL (strA) ΔphoA(pvull) ΔompT</i>	(6)
NN100	SF100, <i>lpp D(uncB-C) zid::Tn10</i>	(7)
FTL10	<i>ΔyidC attB::(araC<sup>+</sup> P<sub>BAD</sub> yidC<sup>+</sup>)</i> ; Kan <sup>R</sup>	(8)

### ***In vivo* complementation**

In *E. coli* strain FTL10 (Table S2) the *yidC* gene is under control of the araBAD promoter/operator. Because YidC is an essential protein in *E. coli*, cell growth of the FTL10 strain is strictly dependent on the presence of arabinose in the medium. The *E. coli* FTL10 was transformed with plasmids pTrc99A ( $\Delta$ YidC), pTrcyidC (wild-type), pTrcTmHis (*T. maritima* TM1461), pTrce61-t396His (YidC61), pTrce321-t225His (YidC321) or pTrce346-t198His (YidC346). Cells were streaked out on LB agar plates containing 100 µg/ml ampicillin and glucose (LB<sup>Amp</sup> agar), and incubated for at least 16 h at 37 °C. To examine cell

growth in liquid culture, at least three single colonies from each of the above-mentioned strains were picked and overnight culture grown at 37 °C in LB medium containing 100 µg/ml ampicillin and 0.5% (wt/vol) arabinose. Next morning, precultures were diluted 100-fold in prewarmed LB medium containing 0.5% (wt/vol) glucose and 50 µM IPTG. The optical density at a wavelength of 660 nm (OD<sub>660</sub>) was measured at different time points until cessation of cell growth.

### **Expression and purification of YidC of *Thermotoga maritima***

The plasmid pTrcTmaHis (Table S2) was transformed into competent *E. coli* NN100 cells using the heat-shock method and plated on LB<sup>Amp</sup> agar. Single colony appeared on LB<sup>Amp</sup> agar incubated at 37°C overnight was inoculated into 2 ml LB<sup>Amp</sup> medium and shaken at 37°C for 7h-8h at 180rpm/min. The 2ml LB<sup>Amp</sup> culture was then transferred into 200ml 2XTY<sup>Amp</sup> and shaken at 37°C for 16-17h at 180rpm/min. The 15ml overnight culture was then inoculated into 1L LB<sup>Amp</sup> for 12L large scale expression. The large bacterial cultures were shaken at 30°C with 200rpm/min. 0.5mM IPTG was added into the culture when the OD<sub>600</sub> reached 0.8-1.0. Cells were harvested after 2h after IPTG induction. Harvested cells were either used directly or frozen at -80°C.

Cells from the 12L cultures were resuspended in cold lysis buffer (50 mM Na-phosphate (pH 8.0), 300 mM NaCl, 10% glycerol, and 10 mM imidazole] by gentle shaking. After cell disruption using a French press (Emulsiflex C3) the lysate was centrifuged for 30 min at 10,000 x *g*. Membranes were pelleted by centrifugation of the supernatant at 100,000 x *g* for 1 h and were resuspended in lysis buffer, potter homogenized, and the protein solubilized with 1% (wt/vol) dodecyl-maltoside (DDM). The supernatant obtained after ultracentrifugation was subjected to a cobalt affinity chromatography carried out using batch method. In average, the solubilized 2.5 g membranes were incubated with 1ml cobalt beads at 4°C with a rotator for at least 1h. The beads were washed with 20 ml of lysis buffer containing 0.05% DDM, followed by 10 ml buffer containing 20 mM imidazole and subsequently with 10 ml buffer containing 40 mM imidazole. The protein was eluted with 5x1ml elution buffer containing 300 mM imidazole. The elution fractions 2 to 4 were pooled and concentrated to 1 ml, desalted on a NAP-10 column, using desalting buffer [(50 mM TrisHCl pH 8.5, 100 mM NaCl, 10% glycerol, and 0.12% decyl maltoside (DM)] and concentrated again to >10 mg/ml by the nanodrop method.

### **Reconstitution and *in vitro* transcription, translation and membrane insertion**

Purified *E. coli* F<sub>0</sub>c and TmYidC were reconstituted into *E. coli* phospholipids (Avanti Polar Lipids, Alabaster, NY) using Bio-beads SM-2 (Bio-Rad) as described (9). Both proteins were reconstituted at protein/lipid ratios of 0.08 and 0.16 (wt/wt).

For *in vitro* synthesis of *E. coli* F<sub>0</sub>c using pET20batpE as DNA template, a cell lysate was prepared essentially as in (4). Coupled transcription, translation and membrane insertion reactions were performed as described previously (10) with DTT present at a final concentration of 4 mM. Each reaction was started by addition of T7 polymerase from Fermentas and the Easytag express protein labeling mix (Perkin Elmer). *In vitro* synthesis and insertion were carried out for 30 min at 37°C in the presence of proteoliposomes. A 10% synthesis control was collected, and the remainder of the mixture was subjected to proteinase K treatment (2 mg/ml, 30 min, and 37°C) in order to remove non-inserted proteins. All consecutive steps were performed as described previously (11).

### Crystallization and data collection

Crystallization trials were set up in sitting drops using a Phoenix liquid handling system (Art Robbins Instruments) and different, commercially available, screens (Nextal). Initial crystals diffracting to approximately 25 Å were obtained from several conditions containing PEG3350 as precipitant. A homemade screen with varying pH and PEG3350 concentration was used for optimization and batch-to-batch variation. Detergents, lipids, heavy atoms and small molecules were tested as additives. Crystals diffracting to high resolution were obtained after 2-7 days at 20°C from sitting drops mixed in a 1:1 ratio with crystallization solution 0.1M Bis-Tris, pH5.5, 0.2M di-ammonium hydrogen citrate with 24% PEG 3350.

For data collection, crystals were mounted on cryoloops directly from crystallization drops and frozen in a cold nitrogen gas stream. Data were recorded at 100 K using a rotation angle of 1.0° per image on a MarCCD detector at the protein beamline X06SA, Swiss Light Source (SLS), Paul Scherrer Institut, Villigen. Datasets were processed with the XDS program (4).

### References

1. J. D. Thompson, T. J. Gibson, D. G. Higgins, Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics* **Chapter 2**, Unit 2 3 (2002).
2. C. L. Nesbo, K. E. Nelson, W. F. Doolittle, Suppressive subtractive hybridization detects extensive genomic diversity in *Thermotoga maritima*. *J Bacteriol* **184**, 4475-4488 (2002).
3. E. Amann, B. Ochs, K. J. Abel, Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69**, 301-315 (1988).
4. M. J. Saller, F. Fusetto, A. J. Driessen, *Bacillus subtilis* SpoIIJ and YqjG function in membrane protein biogenesis. *J Bacteriol* **191**, 6749-6757 (2009).
5. D. Hanahan, Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580 (1983).
6. F. Baneyx, G. Georgiou, *In vivo* degradation of secreted fusion proteins by the *Escherichia coli* outer membrane protease OmpT. *J Bacteriol* **172**, 491-494 (1990).
7. N. Nouwen, M. van der Laan, A. J. Driessen, SecDFyajC is not required for the maintenance of the proton motive force. *FEBS Lett* **508**, 103-106 (2001).



8. K. Hatzixanthis, T. Palmer, F. Sargent, A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol Microbiol* **49**, 1377-1390 (2003).
9. M. van der Laan, E. N. Houben, N. Nouwen, J. Luirink, A. J. Driessen, Reconstitution of Sec-dependent membrane protein insertion: nascent FtsQ interacts with YidC in a SecYEG-dependent manner. *EMBO Rep* **2**, 519-523 (2001).
10. M. C. Jewett, J. R. Swartz, Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnol Bioeng* **86**, 19-26 (2004).
11. C. van der Does, J. de Keyzer, M. van der Laan, A. J. Driessen, Reconstitution of purified bacterial preprotein translocase in liposomes. *Methods Enzymol* **372**, 86-98 (2003).