Supporting Information

The architecture of the OmpC-MlaA complex sheds light on the maintenance of outer membrane lipid asymmetry in *Escherichia coli*

Jiang Yeow[1,2,*], Kang Wei Tan[1,*], Daniel A. Holdbrook[3,*], Zhi-Soon Chong[1], Jan K. Marzinek[3,4], Peter J. Bond[3,4,†], Shu-Sin Chng[1,5,†]

[1]Department of Chemistry, National University of Singapore, Singapore 117543;
[2]National University of Singapore Graduate School for Integrative Sciences and Engineering (NGS), Singapore 117456;
[3]Bioinformatics Institute, Agency for Science, Technology, and Research (A*STAR), Singapore 138671;
[4]Department of Biological Sciences, National University of Singapore, Singapore 117543;
[5]Singapore Center for Environmental Life Sciences Engineering, National University of Singapore (SCELSE-NUS), Singapore 117456

[*]These authors contributed equally to this work.

Table of contents

Supplementary Tables.

Table S1. Bacterial strains used in this study

Strains	Relevant genotypes and characteristics	References
MC4100	*F- araD139 Δ(argF-lac) U169 rpsL150 relA1 flbB5301 ptsF25 deoC1 ptsF25 thi*	(1)
NovaBlue	*endA1 hsdR17 (rK12– mK12+) supE44 thi-1 recA1 gyrA96 relA1 lac F' [proA+ B+ lacIq ZΔM15::Tn10]*	Novagen
BL21(λDE3)	*fhuA2 [lon] ompT gal (λDE3) [dcm] ΔhsdS* *λDE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	Novagen
TKW001	BL21(λDE3) Δ*ompF*::*kan*	This study
CZS010	MC4100 Δ*mlaA*::*kan*	(2)
CZS015	MC4100 Δ*ompC*::*kan*	(2)
NR1216	MC4100 Δ*dsbA*::*kan*	(3)
CZS576	MC4100 Δ*ompC*::*kan*-(*P_{rha}*-*tse2*)	This study
CZS594	MC4100 Δ*ompC*::*ompC*	This study
CZS608	MC4100 Δ*ompC*::*$ompC_{R92A}$*	This study
CZS609	MC4100 Δ*ompC*::*$ompC_{R92L}$*	This study
CZS610	MC4100 Δ*ompC*::*$ompC_{G19W}$*	This study
CZS611	MC4100 Δ*ompC*::*$ompC_{G19W/R92L}$*	This study

Table S2. Plasmids used in this study

Plasmids	Relevant genotypes and characteristics	References
pET22b(+)	pT7lac inducible expression vector, contains N-terminal PelB signal peptide for periplasmic localization; Amp^R	Novagen
pET23/42	pT7 inducible expression vector, contains multiple cloning site of pET42a(+) in pET23a(+) backbone; Amp^R	(4)
pSLC-246	Template plasmid encoding kanamycin resistance gene for positive selection and toxin gene (*tse2*) under the control of rhamnose induceable promoter (P_{rhaB}) for negative selection.	(5)
pSup-BpaRS-6TRN	Encodes an orthogonal tRNA and aminoacyl-tRNA synthetase permitting ribosomal incorporation of *p*Bpa at TAG stop codons	(6)
pKM208	A variation of pKM201 expresses the *lacI* repressor gene that keep expression of *red* and *gam* under tight control prior to IPTG induction	(7)
pACYC184	Low copy cloning vector; Cam^R	(8)
pCDFDuet-1	pT7 inducible expression vector; $Spec^R$	Novagen
pDSW206	Promoter down mutations in -35 and -10 of pTrc99a; Amp^R	(9)
pET23/42-*mlaA-His*	Encodes full length MlaA with C-terminal His8 tag; Amp^R (p-*mlaA-His*)	(2)
pET23/42-*dmlaA-His*	Encodes delipidated version of MlaA (a.a. 19-250) with N-terminal PelB signal peptide (for periplasmic localization) and C-terminal His8 tag; Amp^R (p-*dmlaA-His*)	(2)
pCDF-*mlaA-His*	Encodes full length MlaA with C-terminal His8 tag; $Spec^R$	This study
pCDF-*dmlaA-His*	Encodes delipidated version of MlaA (a.a. 19-250) with N-terminal PelB signal peptide (for periplasmic localization) and C-terminal His8 tag; $Spec^R$	This study
pET22b(+)*dmlaA-His*	Encodes delipidated version of MlaA (a.a. 19-250) with N-terminal PelB signal peptide and C-terminal His6 tag; Amp^R	(2)
pACYC184*ompC*	Encodes full length OmpC under its native promoter; Cam^R	(2)
pDSW206*ompC*	Encodes full length OmpC inducible by *lacI* promoter; Amp^R	This study

Table S3. Primers used in this study.

Primers	Sequence (5' to 3')*
ompC_D7B FP	GAAGTTTACAACAAATAGGGCAACAAATTAGATCTGTACGG
ompC_D7B RP	GATCTAATTTGTTGCCCTATTTGTTGTAAACTTCAGCAGCG
ompC_G8B FP	GAAGTTTACAACAAAGACTAGAACAAATTAGATCTGTACGG
ompC_G8B RP	GATCTAATTTGTTCTAGTCTTTGTTGTAAACTTAGCAGCG
ompC_F40B FP	CCTACATGCGTCTTGGCTAGAAAGGTGAAACTCAGG
ompC_F40B RP	GTTTCACCTTTCTAGCCAAGACGCATGTAGGTCTGG
ompC_L50B FP	G GTT ACT GAC CAGTAGACC GGT TAC GGC CAG TG
ompC_L50B RP	GCC GTA ACC GGTCTACTG GTC AGT AAC CTG AGT TTC
ompC_Y53B FP	CCAGCTGACCGGTTAGGGCCAGTGGGAATATC
ompC_Y53B RP	TCCCACTGGCCCTAACCGGTCAGCTGGTCAGTAAC
ompC_L80B FP	GCA TTC GCA GGTTAGAAA TTC CAG GAT GTG GG
ompC_L80B RP	CATCCTGGAATTTCTAACCTGCGAATGCCACAC
ompC_K81B FP	CATTCGCAGGTCTGTAGTTCCAGGATGTGGGTTC
ompC_K81B RP	CACATCCTGGAACTACAGACCTGCGAATGCCACAC
ompC_F82B FP	GGCATTCGCAGGTCTGAAATAGCAGGATGTGGGTTC
ompC_F82B RP	GTCGAAAGAACCCACATCCTGCTATTTCAGACCTGC
ompC_Q83B FP	GGTCTGAAATTCTAGGATGTGGGTTCTTTCGAC
ompC_Q83B RP	AGAACCCACATCCTAGAATTTCAGACCTGCG
ompC_G86B FP	TTC CAG GAT GTGTAGTCT TTC GAC TAC GGT CGT AAC
ompC_G86B RP	GTAGTCGAAAGACTACACATCCTGGAATTTCAGACCTGC
ompC_F88B FP	GATGTGGGTTCTTAGGACTACGGTCGTAACTACGG
ompC_F88B RP	ACGACCGTAGTCCTAAGAACCCACATCCTG G
ompC_Y90B FP	GGTTCTTTCGACTAGGGTCGTAACTACGGCG
ompC_Y90B RP	GTAGTTACGACCCTAGTCGAAAGAACCCACATCCTG
ompC_A129B FP	GGTAACGGCTTCTAGACCTACCGTAACACTGAC
ompC_A129B RP	GTTACGGTAGGTCTAGAAGCCGTTACCACGCTG
ompC_Y131B FP	CGGCTTCGCGACCTAGCGTAACACTGACTTCTTC
ompC_Y131B RP	GTCAGTGTTACGCTAGGTCGCGAAGCCGTTACC
ompC_N133B FP	GCGACCTACCGTTAGACTGACTTCTTCGGTCTG
ompC_N133B RP	GAAGAAGTCAGTCTAACGGTAGGTCGCGAAGCC
ompC_G138B FP	CACTGACTTCTTCTAGCTGGTTGACGGCCTGAACTTTGC
ompC_G138B RP	GGCCGTCAACCAGCTAGAAGAAGTCAGTGTTACGG
ompC_L143B FP	CTGGTTGACGGCTAGAACTTTGCTGTTCAGTACC
ompC_L143B RP	CAGCAAAGTTCTAGCCGTCAACCAGACCGAAG
ompC_Y149B FP	CTTTGCTGTTCAGTAGCAGGGTAAAAACGGCAAC
ompC_Y149B RP	GTTTTTACCCTGCTACTGAACAGCAAAGTTCAGGCCG
ompC_G151B FP	GTTCAGTACCAGTAGAAAAACGGCAACCCATCTGGTG
ompC_G151B RP	GTTGCCGTTTTTCTACTGGTACTGAACAGCAAAGTTC
ompC_Q266B FP	GTTGCTCAGTACTAGTTCGACTTCGGTCTGCGTC
ompC_Q266B RP	CCGAAGTCGAACTAGTACTGAGCAACAGCTTCG
ompC_F267B FP	GCTCAGTACCAGTAGGACTTCGGTCTGCGTCCG

ompC_F267B RP	CAGACCGAAGTCCTACTGGTACTGAGCAACAGC
ompC_L271B FP	GTTCGACTTCGGTTAGCGTCCGTCCCTGGCTTAC
ompC_L271B RP	CAGGGACGGACGCTAACCGAAGTCGAACTGGTAC
ompC_P273B FP	CTTCGGTCTGCGTTAGTCCCTGGCTTACCTGCAG
ompC_P273B RP	GTAAGCCAGGGACTAACGCAGACCGAAGTCGAACTGG
ompC_L275B FP	GCGTCCGTCCTAGGCTTACCTGCAGTCTAAAG
ompC_L275B RP	GCAGGTAAGCCTAGGACGGACGCAGACCGAAG
ompC_A302B FP	GTTGATGTTGGTTAGACCTACTACTTCAACAAAAACATGTCC
ompC_A302B RP	GAAGTAGTAGGTCTAACCAACATCAACATATTTCAGGATATC
ompC_Y304B FP	GTTGGTGCTACCTAGTACTTCAACAAAAACATGTCC
ompC_Y304B RP	TTTGTTGAAGTACTAGGTAGCACCAACATCAACATATTTCAG
ompC_M310B FP	CTTCAACAAAAACTAGTCCACCTACGTTGACTACAAAATC
ompC_M310B RP	CAACGTAGGTGGACTAGTTTTTGTTGAAGTAGTAGG
ompC_L340B FP	AACATCGTAGCTTAGGGTCTGGTTTACCAGTTC
ompC_L340B RP	GTA AAC CAG ACCCTAAGC TAC GAT GTT ATC AGT GTTG
ompC_NS_N5	ATGAAAGTTAAAGTACTGTCCCTCCTGGTCCCAGCTCTGCGTGTAG GCTGGAGCTGCTTC
ompC_NS_C3	TTAGAACTGGTAAACCAGACCCAGAGCTACGATGTTATCACATATG AA TATCCTCCTTAG
ompC_NS_N5_C	ATGAAAGTTAAAGTACTGTCCCTCCTG
ompC_NS_C3_C	TTAGAACTGGTAAACCAGACCCAG
mlaA_Q126A FP	GAACCCGAAACTGGCGCGGACTGAACCTCACCGC
mlaA_Q126A RP	GGTTCAGTCCGCGCCAGTTTCGGGTTCGCCATC
mlaA_H131A FP	GGACTGAACCTGCGCGCTTCGGTAGTACGCTTG
mlaA_H131A RP	CTACCGAAGCGCGCAGGTTCAGTCCGTTGCAG
mlaA_F152A FP	GTTCAGTTACCGGCGTACGGTAGCTTCACGCTG
mlaA_F152A RP	GAAGCTACCGTACGCCGGTAACTGAACGTAAGG
mlaA_S155A FP	CCGTTCTACGGTGCGTTCACGCTGCGTGATGAC
mlaA_S155A RP	CGCAGCGTGAACGCACCGTAGAACGGTAACTG
mlaA_D160A FP	TTCACGCTGCGTGCGGACGGTGGTGATATGGCG
mlaA_D160A RP	ATCACCACCGTCCGCACGCAGCGTGAAGCTACC
mlaA_D161A FP	CGCTGCGTGATGCGGGTGGTGATATGGCGGATG
mlaA_D161A RP	CATATCACCACCCGCATCACGCAGCGTGAAGC
mlaA_D164A FP	GATGACGGTGGTGCGATGGCGGATGGTTTTTAC
mlaA_D164A RP	ACCATCCGCCATCGCACCACCGTCATCACGCAG
mlaA_D167A FP	GACGGTGGTGATATGGCGGCGGGTTTTTACCCG
mlaA_D167A RP	AAGAACCGGGTAAAAACCCGCCGCCATATCACC
mlaA_V182A FP	GCCGATGTCTGCGGGTAAATGGACGCTTGAAG
mlaA_V182A RP	CGTCCATTTACCCGCAGACATCGGCCAGGTCAG

mlaA_E188A FP	AAATGGACGCTTGCGGGGATCGAAACCCGCGC
mlaA_E188A RP	GTTTCGATCCCCGCAAGCGTCCATTTACCCAC
mlaA_T192A FP	GAAGGGATCGAAGCGCGCGCTCAGCTGCTG
mlaA_T192A RP	CTGAGCGCGCGCTTCGATCCCTTCAAGCGTC
mlaA_Q195A FP	GAAACCCGCGCTGCGCTGCTGGATTCCGATGG
mlaA_Q195A RP	GAATCCAGCAGCGCAGCGCGGGTTTCGATCCC
mlaA_N226A FP	GATTTCATCGCTGCGGGCGGCGAACTCAAACCG
mlaA_N226A RP	GAGTTCGCCGCCCGCAGCGATGAAATCATGACG
mlaA_D61R FP	GTCGCCTGGCGTCGCTATGTTCCGCAACCGGCG
mlaA_D61R RP	TTGCGGAACATAGCGACGCCAGGCGACAGCGAC
mlaA_D160R FP	TTCACGCTGCGTCGCGACGGTGGTGATATGGCG
mlaA_D160R RP	ATCACCACCGTCGCGACGCAGCGTGAAGCTACC
mlaA_D161R FP	CGCTGCGTGATCGCGGTGGTGATATGGCGGATG
mlaA_D161R RP	CATATCACCACCGCGATCACGCAGCGTGAAGC
mlaA_D164R FP	GATGACGGTGGTCGCATGGCGGATGGTTTTTAC
mlaA_D164R RP	ACCATCCGCCATGCGACCACCGTCATCACGCAG
mlaA_D167R FP	GACGGTGGTGATATGGCGCGCGGTTTTTACCCG
mlaA_D167R RP	AAGAACCGGGTAAAAACCGCGCGCCATATCACC
mlaA_E188R FP	AAATGGACGCTTCGCGGGATCGAAACCCGCGC
mlaA_E188R RP	GTTTCGATCCCGCGAAGCGTCCATTTACCCAC
mlaA_D160R D161R FP	TTCACGCTGCGTCGCCGCGGTGGTGATATGGCG
mlaA_D160R D161R RP	ATCACCACCGCGGCGACGCAGCGTGAAGCTACC
mlaA_D160R D164R FP	TTCACGCTGCGTCGCGACGGTGGTCGCATG GCG
mlaA_D160R D164R RP	GCGACCACCGTCGCGACGCAGCGTGAAGCTACC
mlaA_D161R D164R FP	CGCTGCGTGATCGCGGTGGTCGCATGGCGGATG
mlaA_D161R D164R RP	CATGCGACCACCGCGATCACGCAGCGTGAAGC
mlaA_D161R D167R FP	CGCTGCGTGATCGCGGTGGTGATATGGCGCGCG
mlaA_D161R D167R RP	CATATCACCACCGCGATCACGCAGCGTGAAGC
mlaA_D164R D167R FP	GATGACGGTGGTCGCATG GCGCGCGGTTTTTAC
mlaA_D164R D167R RP	ACCGCGCGCCATGCGACCACCGTCATCACGCAG
3D3R FP SDM	CTTCACGCTGCGTCGCCGCGGTGGTCGC ATGGCGGATGGTTTTTACC
3D3R RP SDM	AACCATCCGCCATGCGACCACCGCGGCGACGCAGCGTGA AGCTACCG
F^{152}YGSF_to_5A FP	GTTCAGTTACCGGCGGCGGCGGCGGCGACGCTGCGTGATG ACGGTGG
F^{152}YGSF_to_5A RP	CATCACGCAGCGTCGCCGCCGCCGCCGCCGGTAACTGAAC GTAAGG
GVGYG_3G3A_FP	CTTGGTCATTATGCGGTGGCGTATGCGCCTTACGTTCAGTT ACCG
GVGYG_3G3A_RP	CTGAACGTAAGGCGCATACGCCACCGCATAATGACCAAG CGTAC
GVGYG_3G3P_FP	CTTGGTCATTATCCTGTGCCTTATCCTCCTTACGTTCAGTT ACCG

GVGYG_3G3P_RP	CTGAACGTAAGGAGGATAAGGCACAGGATAATGACCAAG CGTAC
mlaA_P151A FP	TACGTTCAGTTAGCGTTCTACGGTAGCTTCACGCTG
mlaA_P151A RP	GCTACCGTAGAACGCTAACTGAACGTAAGGCCC
Y^{147}VQL_to_4A FP	GGTTATGGGCCTGCGGCGGCGGCGCCGTTCTACGGTAGCT TCAC
Y^{147}VQL_to_4A RP	CTACCGTAGAACGGCGCCGCCGCCGCAGGCCCATAACCCA CGCC
mlaA_M39C FP	TTCAACCGCACCTGCTACAACTTCAACTTCAATG
mlaA_M39C RP	AGTTGAAGTTGTAGCAGGTGCGGTTGAACCCTTC
mlaA_Y40C FP	CAACCGCACCATGTGCAACTTCAACTTCAATG
mlaA_Y40C RP	AGTTGAAGTTGCACATGGTGCGGTTGAACCC
mlaA_F42C FP	ACCATGTACAACTGCAACTTCAATGTATTAGAC
mlaA_F42C RP	TACATTGAAGTTGCAGTTGTACATGGTGCGGTT
mlaA_N43C FP	CATGTACAACTTCTGCTTCAATGTATTAGACCCG
mlaA_N43C RP	TAATACATTGAAGCAGAAGTTGTACATGGTGCGG
mlaA_D48C FP	CTTCAATGTATTATGCCCGTATATTGTTCGACC
mlaA_D48C RP	ACAATATACGGGCATAATACATTGAAGTTGAAG
mlaA_D61C FP	GTCGCCTGGCGTTGCTATGTTCCGCAACCGGCG
mlaA_D61C RP	TTGCGGAACATAGCAACGCCAGGCGACAGCGAC
mlaA_F74C FP	GTTTGAGCAACTGCACTGGCAACCTTGAAGAACC
mlaA_F74C RP	CAAGGTTGCCAGTGCAGTTGCTCAAACCGTTACG
mlaA_L78C FP	CTTTACTGGCAACTGCGAAGAACCTGCGGTGATGG
mlaA_L78C RP	CGCAGGTTCTTCGCAGTTGCCAGTAAAGTTGCTCAAAC
mlaA_M84C FP	GAACCTGCGGTGTGCGTTAACTACTTCTTGCAGG
mlaA_M84C RP	GAAGTAGTTAACGCACACCGCAGGTTCTTCAAGG
mlaA_N86C FP	GCGGTGATGGTTTGCTACTTCTTGCAGGGCGA
mlaA_N86C RP	CTGCAAGAAGTAGCAAACCATCACCGCAGGTTCTTC
mlaA_D92C FP	TAACTACTTCTTGCAGGGCTGCCCTTATCAGGGG
mlaA_D92C RP	GACCATCCCCTGATAAGGGCAGCCCTGCAAGAAG
mlaA_T107C FP	CGCTTTTTCCTGAACTGCATTTTGGGGATGGGCGG
mlaA_T107C RP	CATCCCCAAAATGCAGTTCAGGAAAAAGCGGGTAAAGTG G
mlaA_F114C FP	GGGATGGGCGGTTGCATTGATGTTGCAGGGATG
mlaA_F114C RP	GCAACATCAATGCAACCGCCCATCCCCAAAATGG
mlaA_Q126C FP	GAACCCGAAACTGTGCCGGACTGAACCTCACCGC
mlaA_Q126C RP	GGTTCAGTCCGGCACAGTTTCGGGTTCGCCATC
mlaA_T136C FP	CGCTTCGGTAGTTGCCTTGGTCATTATGGCGTG
mlaA_T136C RP	ATAATGACCAAGGCAACTACCGAAGCGGTGAGG
mlaA_Y144C FP	ATGGCGTGGGTTGCGGGCCTTACGTTCAGTTACC
mlaA_Y144C RP	GAACGTAAGGCCCGCAACCCACGCCATAATGAC
mlaA_Q149C FP	GGGCCTTACGTTTGCTTACCGTTCTACGGTAGC
mlaA_Q149C RP	GTAGAACGGTAAGCAAACGTAAGGCCCATAACC
mlaA_L150C FP	TTACGTTCAGTGCCCGTTCTACGGTAGCTTC
mlaA_L150C RP	ACCGTAGAACGGGCACTGAACGTAAGGCCC

mlaA_T157C FP	CTACGGTAGCTTCTGCCTGCGTGATGACGGTGG
mlaA_T157C RP	TCATCACGCAG GCAGAAGCTACCGTAGAACGG
mlaA_D161C FP	CGCTGCGTGATTGCGGTGGTGATATGGCGGATG
mlaA_D161C RP	CATATCACCACCGCAATCACGCAGCGTGAAGC
mlaA_D167C FP	GACGGTGGTGATATGGCGTGCGGTTTTTACCCG
mlaA_D167C RP	AAGAACCGGGTAAAAACCGCACGCCATATCACC
mlaA_K184C FP	CCGATGTCTGTGGGTTGCTGGACGCTTGAAG
mlaA_K184C RP	GATCCCTTCAAGCGTCCAGCAACCCACAGAC
mlaA_E188C FP	AAATGGACGCTTTGCGGGATCGAAACCCGCGC
mlaA_E188C RP	GTTTCGATCCCGCAAAGCGTCCATTTACCCAC
mlaA_T192C FP	GAAGGGATCGAATGCCGCGCTCAGCTGCTG
mlaA_T192C RP	CTGAGCGCGGCATTCGATCCCTTCAAGCGTC
mlaA_R193C FP	GGGATCGAAACCTGCGCTCAGCTGCTGGATTCC
mlaA_R193C RP	CAGCAGCTGAGCGCAGGTTTCGATCCCTTCAAGC
mlaA_Q195C FP	GAAACCCGCGCTTGCCTGCTGGATTCCGATGG
mlaA_Q195C RP	GAATCCAGCAGGCAAGCGCGGGTTTCGATCCC
mlaA_R204C FP	GATTCCGATGGTCTGCTGTGCCAGTCGTCCGATCC
mlaA_R204C RP	AATATAAGGATCGGACGACTGGCACAGCAGACCATC
mlaA_P209C FP	CAGTCGTCCGATTGCTATATTATGGTGCGCGAAG
mlaA_P209C RP	GCACCATAATATAGCAATCGGACGACTGACGCAG
mlaA_R220C FP	GCGAAGCGTACTTCCAGTGCCATGATTTCATC
mlaA_R220C RP	CATTAGCGATGAAATCATGGCACTGGAAGTAC
mlaA_N226C FP	GATTTCATCGCTTGCGGCGGCGAACTCAAACCG
mlaA_N226C RP	GAGTTCGCCGCCGCAAGCGATGAAATCATGACG
mlaA_G227C FP	TTCATCGCTAATTGCGGCGAACTCAAACCGCAG
mlaA_G227C RP	GTTTGAGTTCGCCGCAATTAGCGATGAAATCATG
pCDFDuet-1_pelB_mlaA_ChisNdeI_Fwd	CGCTCATATGAAATACCTGCTGCCGACCGCTGCTGC
pCDFDuet-1_FL_mlaA_ChisNdeI_Fwd	CGCTCATATGAAGCTTCGCCTGTCG
pCDFDuet-1_mlaA_AvrII Rev	AGATCCTAGGTCAGTGGTGGTGGTGGTGGTGCTCGAG
pDSW206_ompC_NcoI Fwd	CGATCCATGGCAAAAGTTAAAGTACTGTCCCTCC
pDSW206_ompC_HindIII Rev	CGCTAAGCTTTTAGAACTGGTAAACCAGACCCAGAGC

* sites for mutagenesis or restriction enzyme cleavage, where relevant, are underlined.

Table S4. Summary of all-atom molecular simulations: system compositions and simulation times

Protein Configuration	Lipids	Water and Ions	Simulation time (# of simulations x ns)
MlaA	N/A	9439 H_2O 29 K^+ 19 Cl^-	1 x 500
MlaA	272 DMPE	11734 H_2O 42 K^+ 32 Cl-	1 x 500
OmpC trimer MlaA (ClusPro model) in orientation 1	980 DMPE	36113 H_2O 98 K^+ 98 Cl^-	1 x 500 1 x 320 1 x 130
OmpC trimer MlaA (ClusPro model) in orientation 2	980 DMPE	36113 H_2O 98 K^+ 98 Cl^-	1 x 500 1 x 500

Supplementary Figures

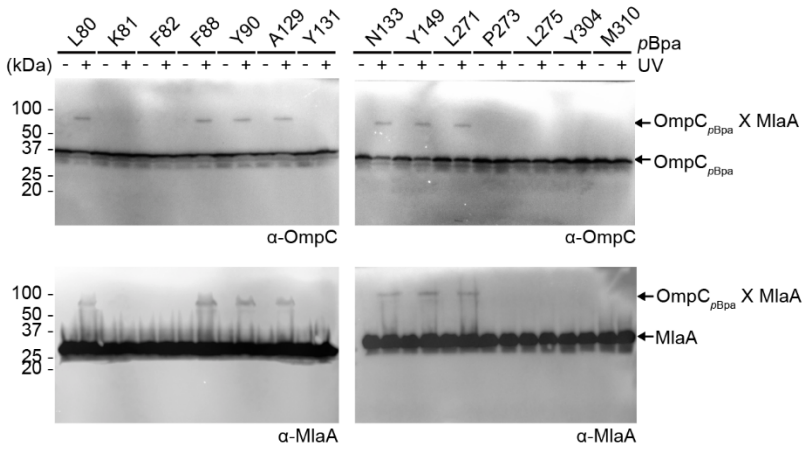

Figure S1. Seven more positions at the dimeric interface of the OmpC trimer contact MlaA. Immunoblots showing UV-dependent formation of crosslinks between OmpC and MlaA in Δ*ompC* cells expressing OmpC substituted with *p*Bpa at indicated positions, selected as part of the localized search.

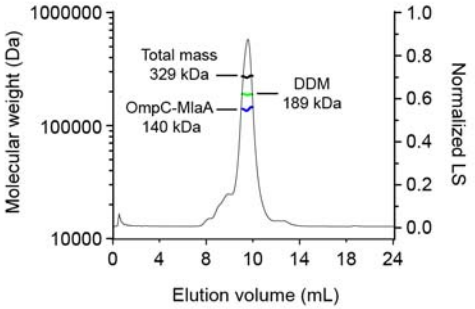

Figure S2. SEC-MALS analysis of the OmpC-MlaA complex revealing that one copy of MlaA binds to the OmpC trimer. As indicated, total molecular mass: 329 (± 0.4%) kDa; protein molecular mass: 140 (± 0.4%) kDa (observed), 148 kDa (predicted, $OmpC_3MlaA$); modifier (DDM) molecular mass: 189 (± 0.8%) kDa. Numbers stated after ± show statistical consistency of analysis.

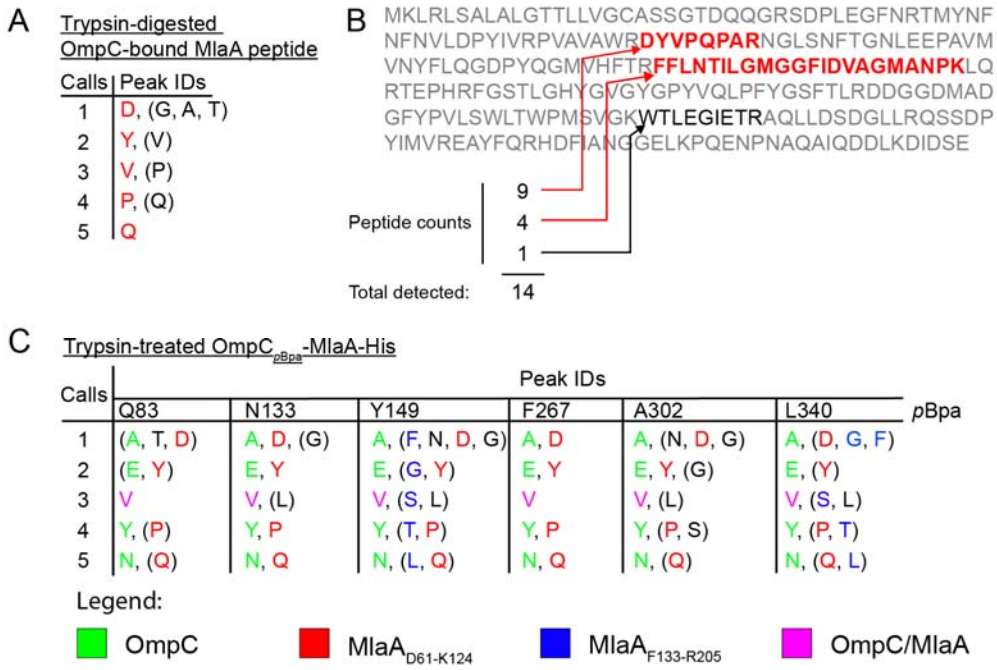

Figure S3. N-terminal sequencing and MS/MS analyses identified two specific MlaA peptides binding to OmpC. (*A*) First five residue calls for the MlaA peptide remaining bound to OmpC after trypsin digestion (see Fig. 2*A*) revealed that it starts with D^{61}YVPQ of full-length MlaA protein. (*B*) MS/MS analysis of the MlaA peptide remaining bound to OmpC after trypsin digestion detected two MlaA fragments with high peptide counts (sequences colored *red*), suggesting that the OmpC-bound peptide has boundaries from D61 to K124. (*C*) First five residue calls for protein bands containing MlaA peptides crosslinked to OmpC$_{p\text{Bpa}}$ (see Fig. 2*B*) revealed the presence of MlaA peptides starting with D^{61}YVPQ and F^{133}GSTL, along with OmpC N-terminus A^{21}EVYN. Residue calls are assigned to the respective protein/peptide as denoted by the legend.

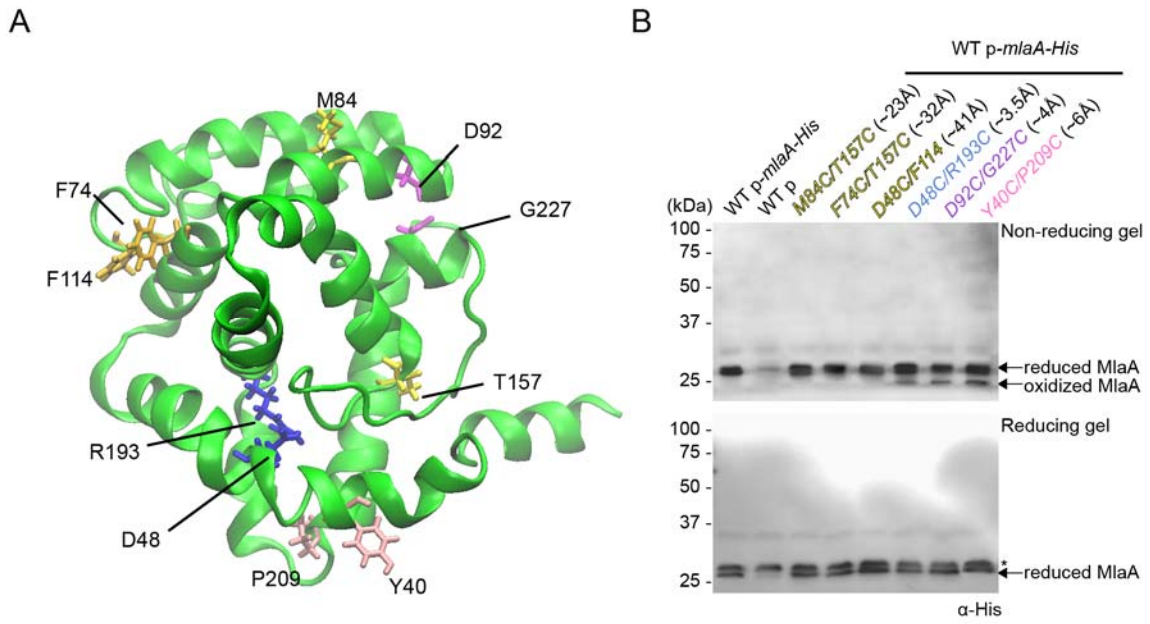

Figure S4. Residue pairs on MlaA predicted to contact each other based on coevolution analysis allow the formation of disulfide bonds when substituted with cysteines. (*A*) Cartoon representation of the MlaA structural model predicted based on residue-residue contacts inferred from co-evolution analysis of metagenomic sequence data prediction (GREMLIN, (10)), with strongly co-evolved residue pairs that are mutated to cysteines highlighted (same colored sticks). The figure was generated using the program PyMOL (12). (*B*) Immunoblots showing oxidized or reduced forms of indicated MlaA-His double cysteine variants expressed in wild-type cells from the pET23/42 vector (p). Samples were subjected to non-reducing (*top*) or reducing (*bottom*) SDS-PAGE prior to transfer. A protein that cross-reacted with the α-His antibody is denoted with (*). Distances between cysteine pairs in unit angstrom (Å), as measured in the model in (*A*), are indicated in parentheses.

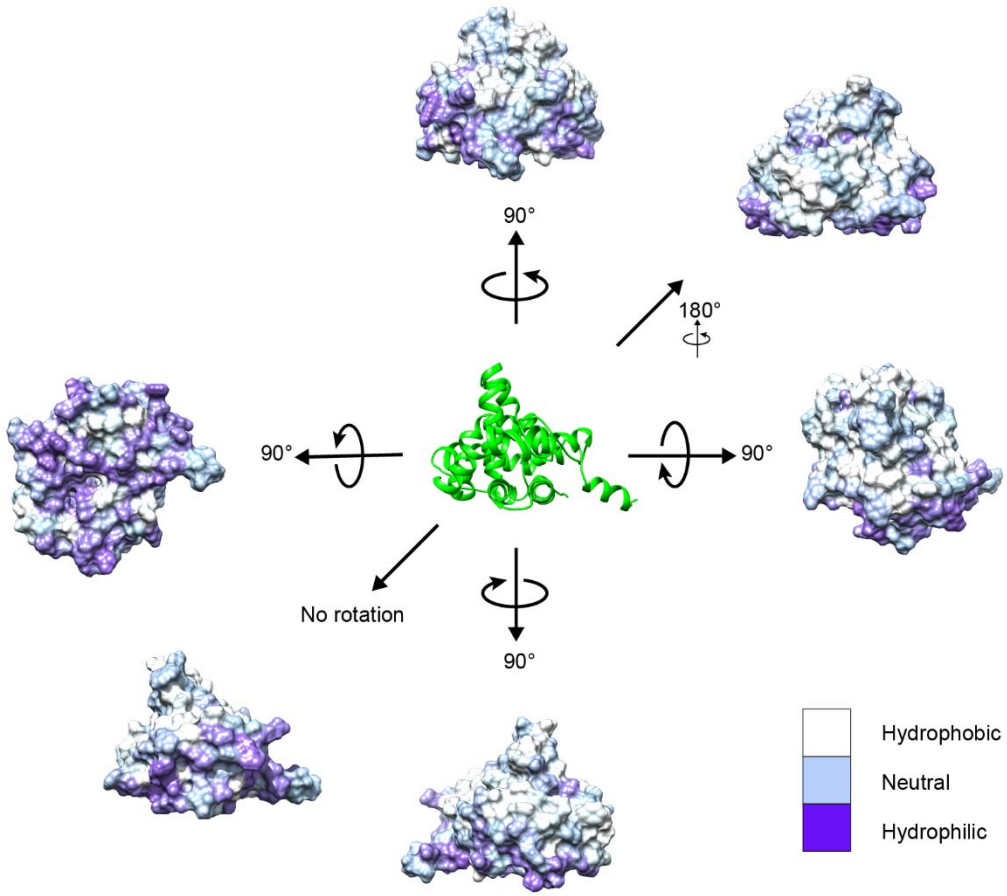

Figure S5. The surface of MlaA is mostly hydrophobic. Surface representation of the MlaA model (10) depicted in multiple orientations and colored based on amino acid hydrophobicity. Purple, light blue and white represent most hydrophilic to most hydrophobic amino acids based on the Kyte-Doolittle scale (11). The figures were generated using the program Chimera (13).

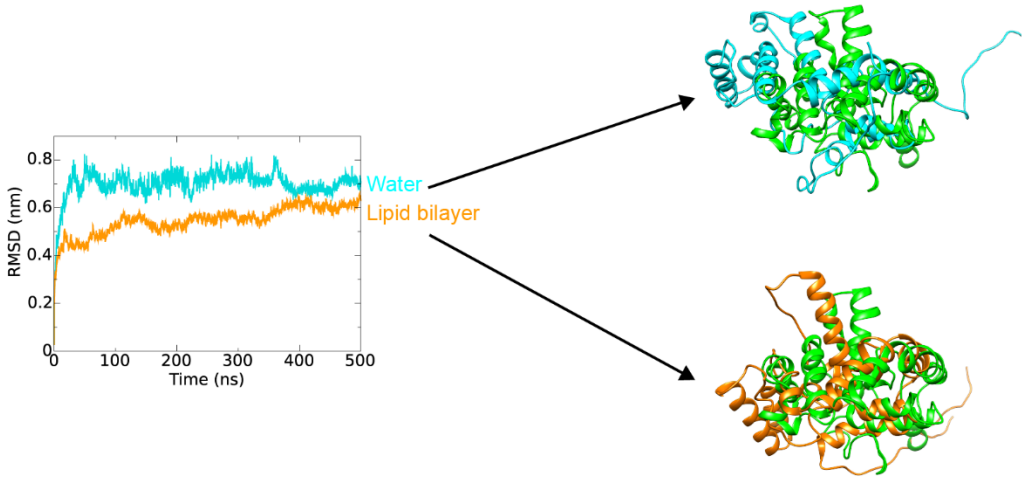

Figure S6. The MlaA structure modelled from co-evolution analysis (10) is more stable in the lipid bilayer. Averaged root-mean-square-deviation (RMSD) plots illustrating the changes of the backbone of MlaA models over the course of all-atomistic MD simulations, when placed in water (*cyan*) or in a lipid bilayer (*orange*). Superimpositions of the initial (*green*) and final structures for each simulation are shown on the right. The figures were generated using the program Chimera (13).

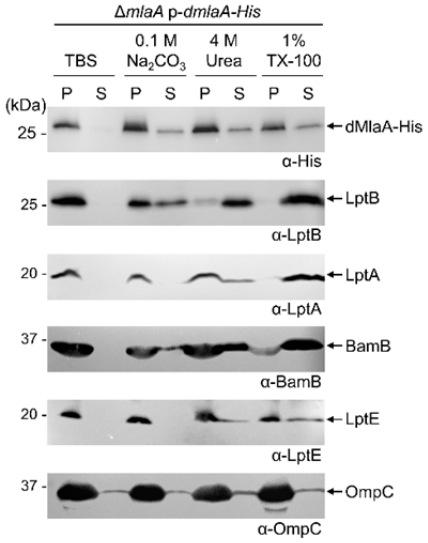

Figure S7. MlaA behaves like an integral membrane protein and is resistant to extraction from membranes under various conditions. Immunoblots showing extraction profiles of delipidated MlaA-His (dMlaA-His) from total membranes upon incubation with high pH (0.1 M Na_2CO_3), chaotropic (4 M urea), or mild detergent (1% (v/v) TX-100) solutions for 1 hour. Samples were subjected to immunoblot analyses after fractionation (insoluble membrane pellet (P) and soluble (S) fractions) by centrifugation. Known peripheral membrane proteins (LptA and LptB), OM lipoproteins (BamB and LptE), and β-barrel proteins (OmpC) are used as controls. Even though both LptA and LptB are peripheral membrane proteins, they exhibit different membrane extraction profiles; while both proteins are easily extracted by 1% TX-100, LptA is more resistant to extraction by 4 M urea. The two OM lipoproteins also exhibit different membrane extraction profiles. Being a lipoprotein with its protein domain residing entirely in the periplasm, BamB appears to have an extraction profile similar to LptA. In contrast, LptE, which is embedded within the lumen of the LptD β-barrel domain, behaves like an integral membrane protein, such as OmpC, and is essentially not extracted from the membrane under the various conditions.

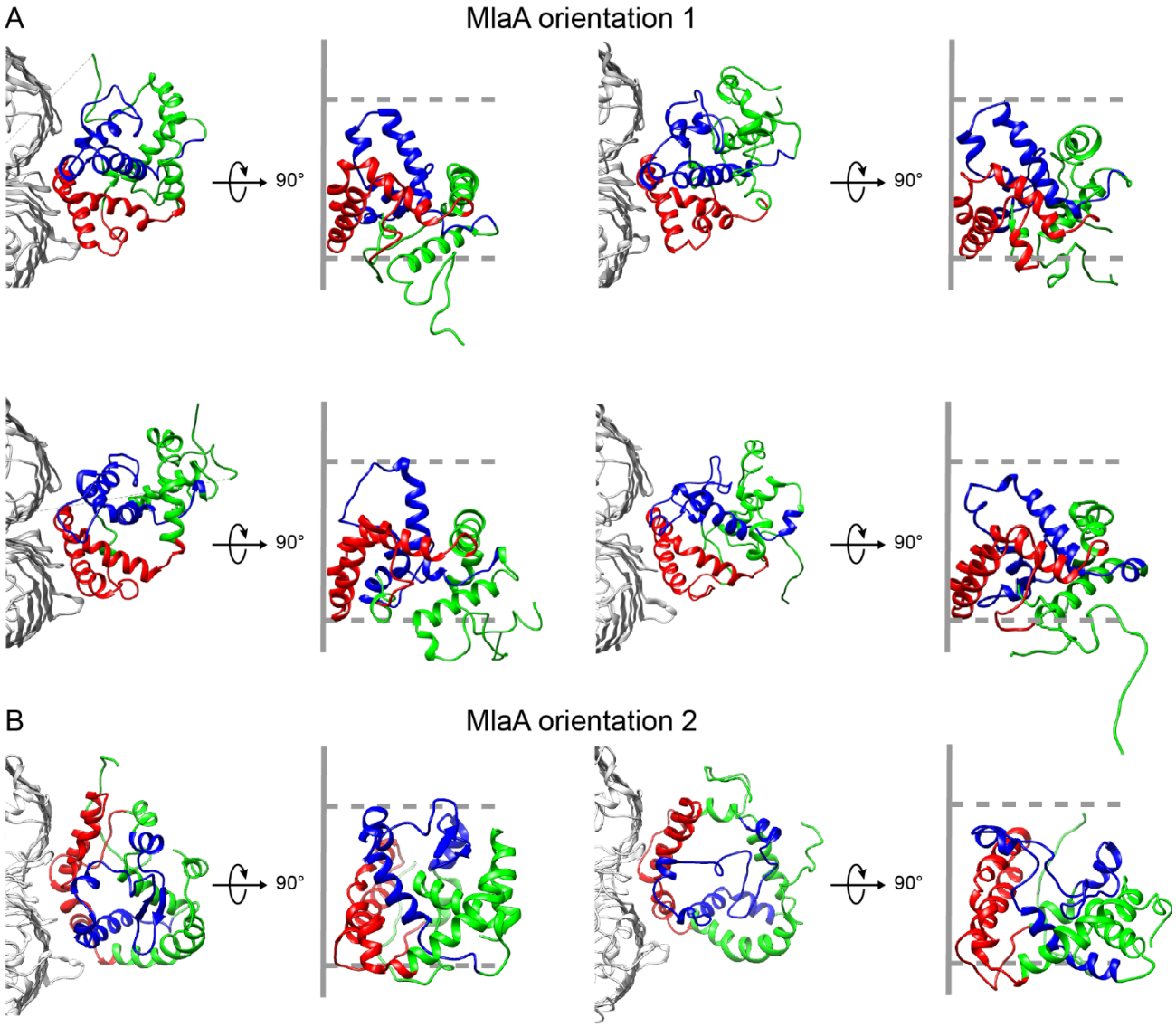

Figure S8. Six major clusters of all-atomistic MD simulated OmpC-MlaA structure depict how MlaA interacts with OmpC in two possible orientations in the OM bilayer. The bottom right model in (*A*) and (*B*) are reproduced as representative models in Figs. 3*A* and 3*B*. $MlaA_{D61\text{-}K124}$ and $MlaA_{F133\text{-}R205}$ peptides are highlighted in *red* and *blue*, respectively, as in Fig. 2*D*. The OM boundaries are indicated as *gray* dashed lines. The figures were generated using the program Chimera (13).

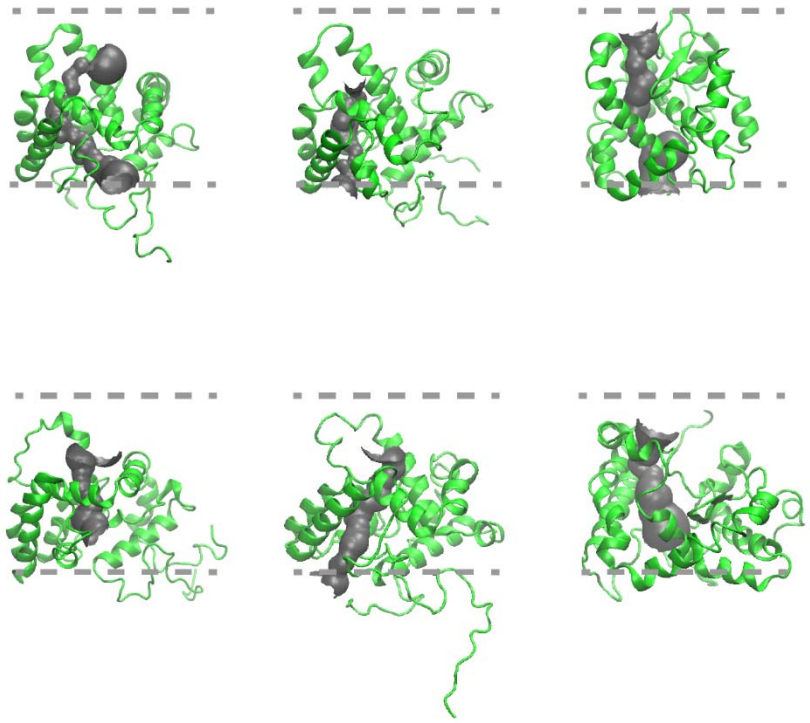

Figure S9. All six major clusters of MlaA structure from all-atomistic MD simulations of the OmpC-MlaA complex with putative hydrophilic channels depicted in *gray*. The bottom right model is reproduced in Fig. 4*A*. The OM boundaries are indicated as *gray* dashed lines. The figures were generated using the program VMD (14).

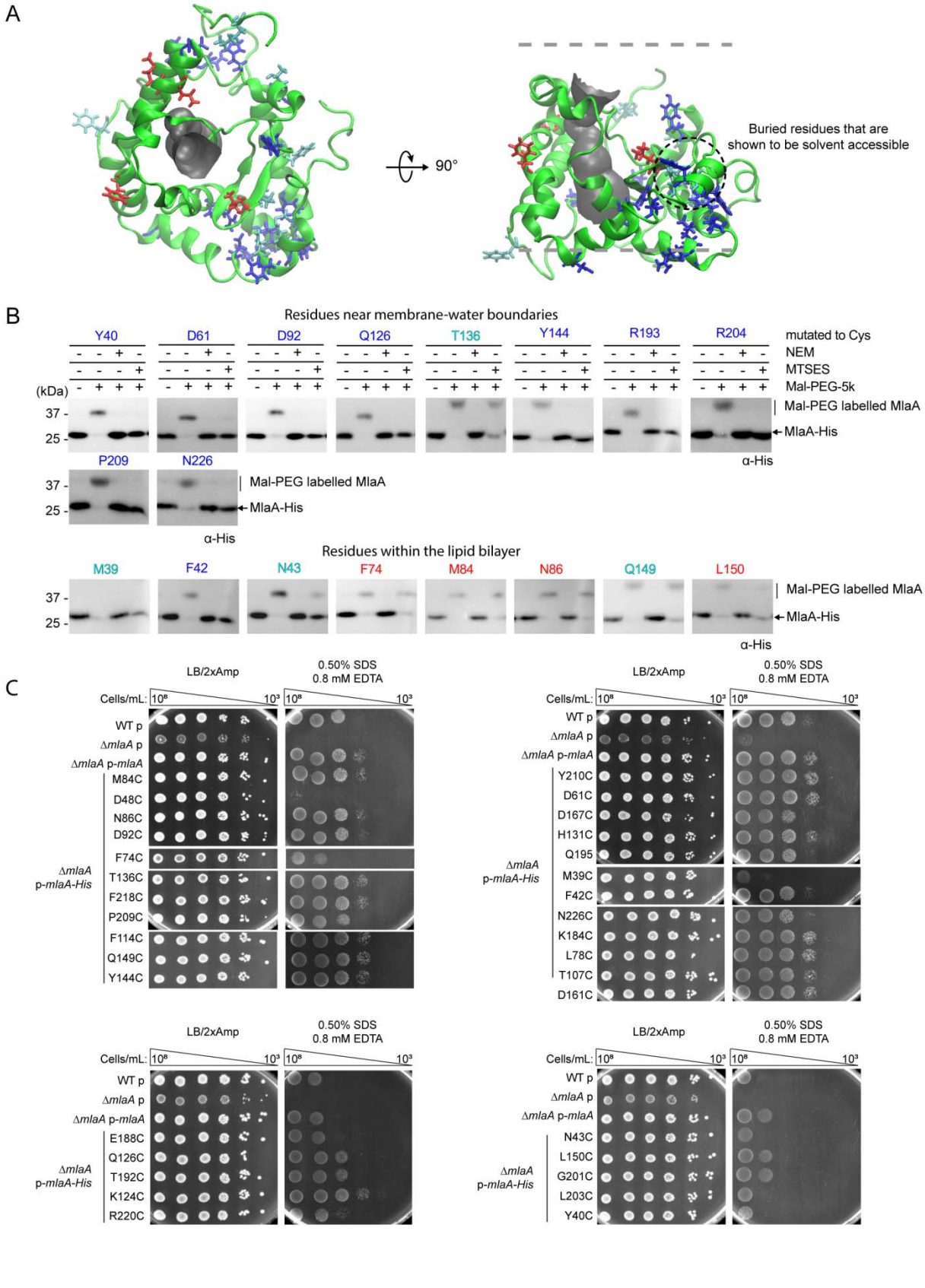
A
90°
Buried residues that are shown to be solvent accessible
B
Residues near membrane-water boundaries
Y40
D61
D92
Q126
T136
Y144
R193
R204
mutated to Cys
NEM
MTSES
Mal-PEG-5k
(kDa)
37
25
Mal-PEG labelled MlaA
MlaA-His
α-His
P209
N226
Residues within the lipid bilayer
M39
F42
N43
F74
M84
N86
Q149
L150
C
LB/2xAmp
0.50% SDS
0.8 mM EDTA
Cells/mL:
WT p
ΔmlaA p
ΔmlaA p-mlaA
ΔmlaA p-mlaA-His
M84C
D48C
N86C
D92C
F74C
T136C
F218C
P209C
F114C
Q149C
Y144C
Y210C
D61C
D167C
H131C
Q195
M39C
F42C
N226C
K184C
L78C
T107C
D161C
E188C
Q126C
T192C
K124C
R220C
N43C
L150C
G201C
L203C
Y40C

Figure S10. Substituted cysteine accessibility for residues in MlaA largely agrees with their predicted locations (near/at membrane-water boundaries or buried within the lipid bilayer). (*A*) A representative structure of MlaA from all-atomistic MD simulations with its putative channel depicted in *gray*. Non-channel residues that are fully, partially, or not solvent accessible, based on SCAM in (*B*), are highlighted in *blue*, *cyan*, and *red*, respectively. The figures were generated using the program VMD (14). (*B*) Immunoblots showing maleimide-polyethylene glycol (Mal-PEG) alkylation of MlaA variants containing channel-facing residues substituted with cysteine (as depicted in (*A*)) following labelling by membrane permeable *N*-ethylmaleimide (NEM) or impermeable (MTSES) reagents. Mal-PEG alkylated $MlaA_{Cys}$-His variants show a ~5 kDa mass shift. Positions fully, partially, or not blocked by MTSES, which reflects the level of solvent accessibility, are highlighted in *blue*, *cyan*, or *red*, respectively. (*C*) Analysis of SDS/EDTA sensitivity of wild-type (WT) and Δ*mlaA* strains producing indicated MlaA cysteine variants from the pET23/42 vector (p).

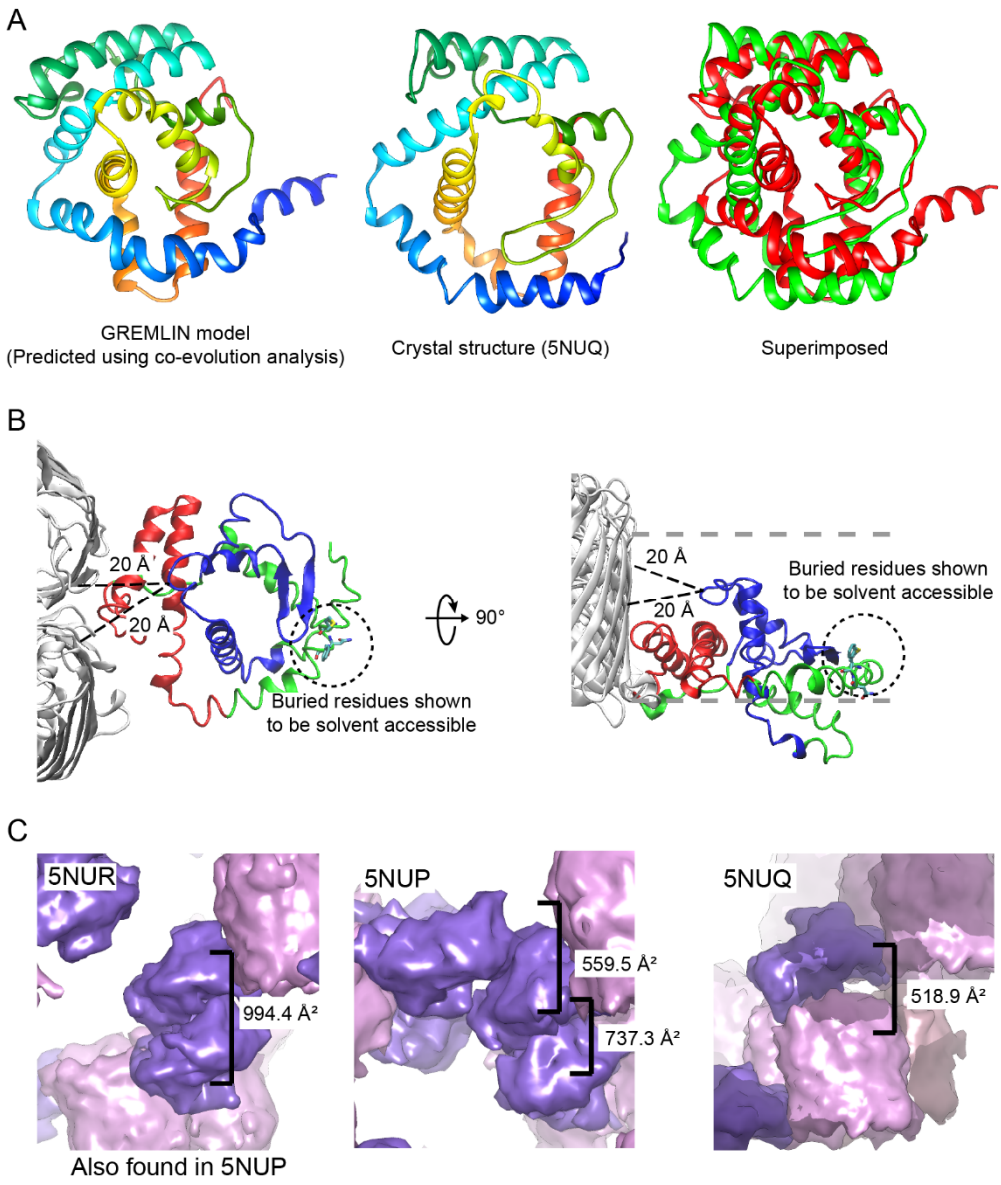

Figure S11. Brief analyses of the crystal structures of MlaA-porin complexes. (*A*) Side-by-side comparison of MlaA model predicted by co-evolution analysis (*left*) with the crystal structure of MlaA derived from the OmpF-MlaA complex (PDB ID: 5NUQ; *middle*). A superimposition of these structures is shown on the right. (*B*) Cartoon representation of the OmpF-MlaA complex (PDB ID: 5NUQ) in top and side views, with $MlaA_{D61\text{-}K124}$ and $MlaA_{F133\text{-}R205}$ peptides highlighted in *red* and *blue*, respectively (as in Fig. 2*D*). The smallest distances between the $MlaA_{F133\text{-}R205}$ peptide (*blue*) and porin residues equivalent to L149/L340 in *E. coli* OmpC are indicated. MlaA residues presumably buried in the lipid bilayer but solvent accessible (SCAM; Fig. S10*B*) are circled and depicted in sticks. (*C*) Surface representations of MlaA-porin crystal structures illustrating artificial crystal contacts (MlaA-MlaA or MlaA-porin) observed in different crystal forms. The buried surface areas (Å^2) of these contacts are indicated. Porins and MlaA are shown in *plum* and *medium purple*, respectively. All figures were generated using the program Chimera (13).

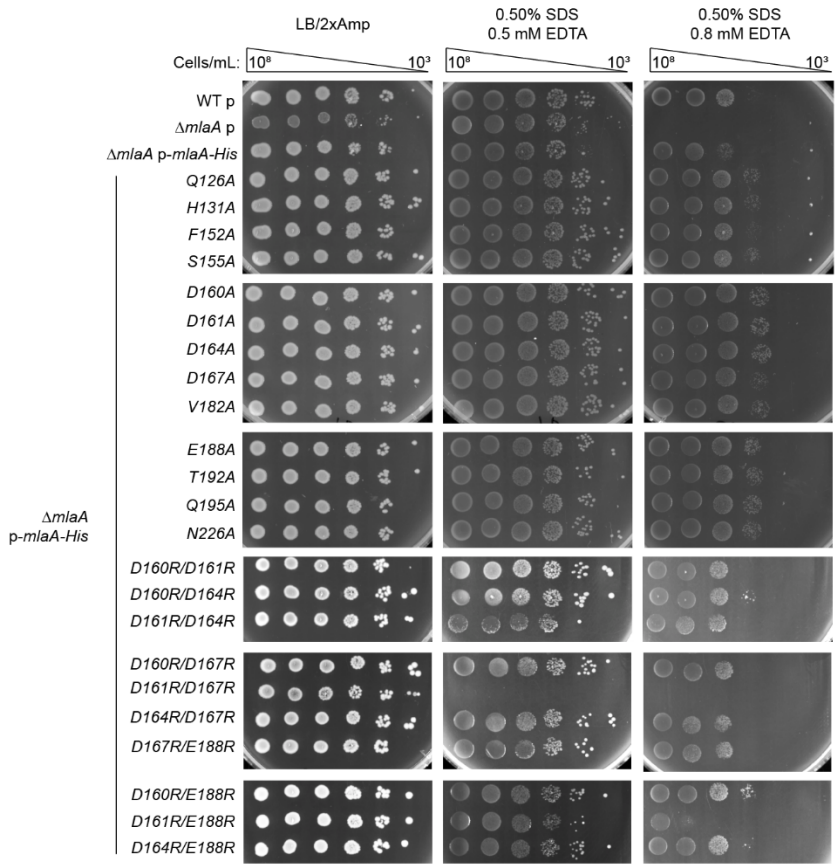

Figure S12. All single alanine mutations and most double arginine substitutions in the channel, except D161R/D167R, do not disrupt function in MlaA. Analysis of SDS/EDTA sensitivity of wild-type (WT) and Δ*mlaA* strains producing indicated MlaA channel variants from the pET23/42 vector (p).

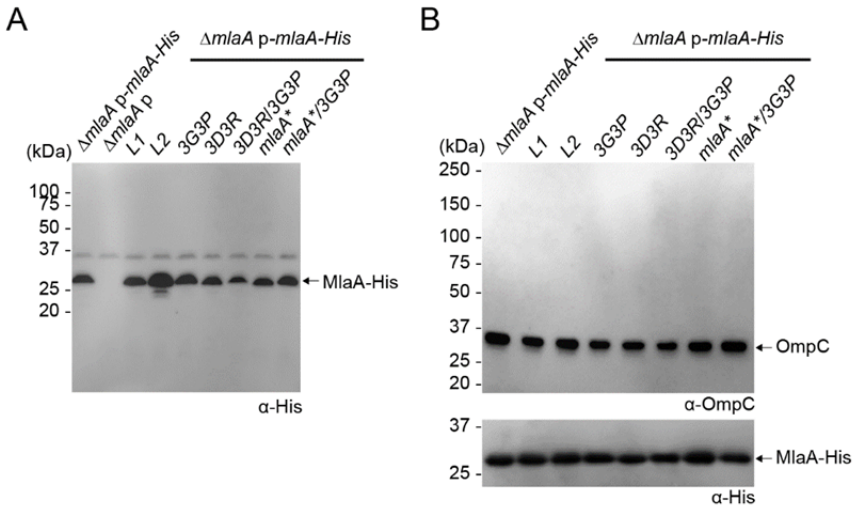

Figure S13. Mutations in functional regions of MlaA do not significantly affect protein levels or its interaction with OmpC. (*A*) Immunoblot showing the levels of indicated MlaA-His variants produced from the pET23/42 vector (p) in the Δ*mlaA* strain. (*B*) Immunoblots showing OmpC copurified with indicated MlaA-His variants produced from the pET23/42 vector (p) in the Δ*mlaA* strain.

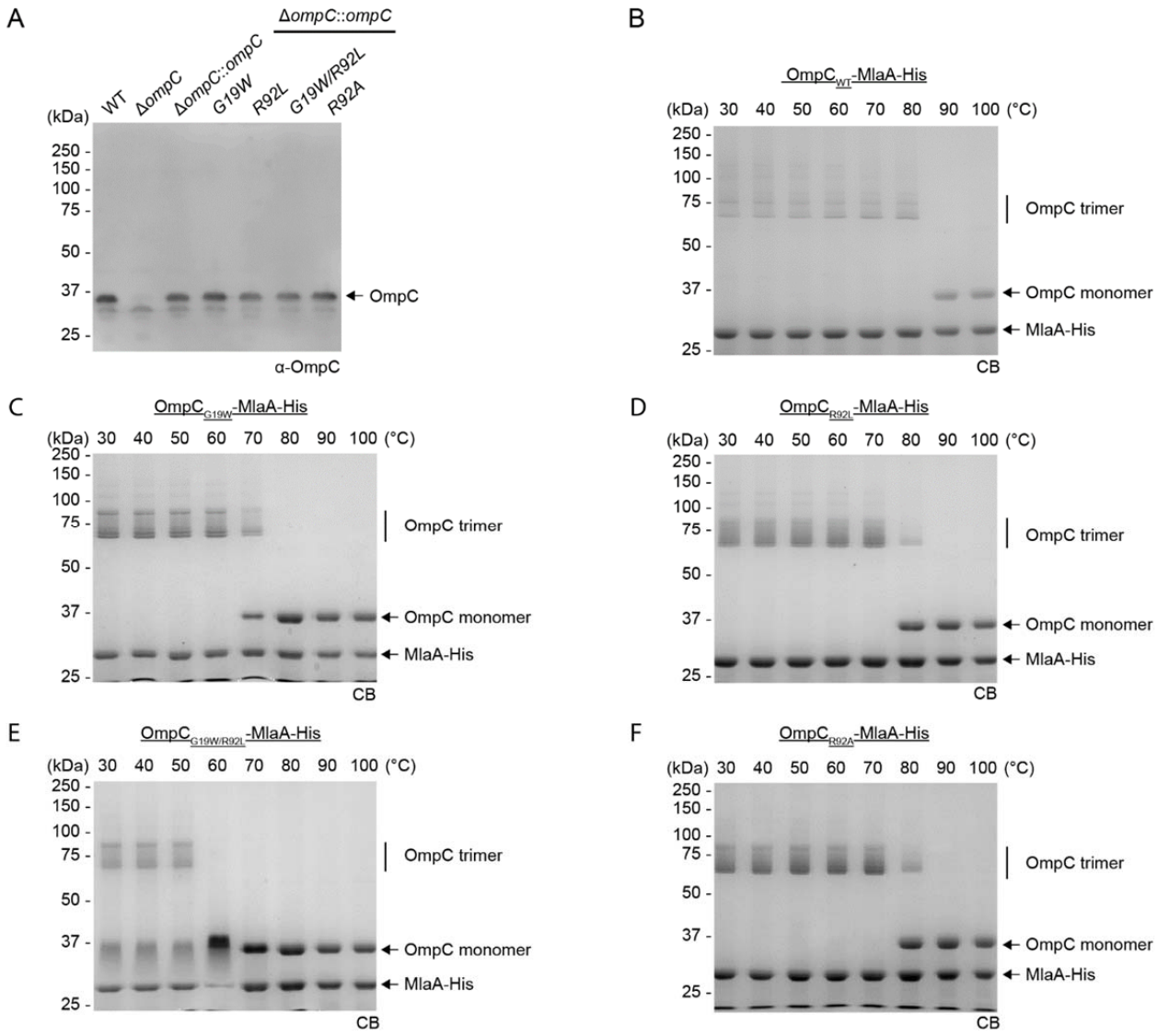

Figure S14. Mutations on residues G19 and R92 do not affect OmpC levels in cells, but weaken trimer stability in vitro. (*A*) Immunoblot showing the levels of wild-type OmpC and indicated OmpC variants produced from the chromosomal locus. (*B-F*) In vitro temperature titration of purified OmpC-MlaA-His and the indicated variants subjected to seminative SDS-PAGE (12% Tris.HCl gel), followed by Coomassie blue (CB) staining.

Supplementary References

1. Casadaban, M. J. (1976) Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. *J Mol Biol* **104**, 541-555
2. Chong, Z. S., Woo, W. F., and Chng, S. S. (2015) Osmoporin OmpC forms a complex with MlaA to maintain outer membrane lipid asymmetry in Escherichia coli. *Mol Microbiol* **98**, 1133-1146
3. Ruiz, N., Chng, S. S., Hiniker, A., Kahne, D., and Silhavy, T. J. (2010) Nonconsecutive disulfide bond formation in an essential integral outer membrane protein. *Proc Natl Acad Sci U S A* **107**, 12245-12250
4. Wu, T., McCandlish, A. C., Gronenberg, L. S., Chng, S. S., Silhavy, T. J., and Kahne, D. (2006) Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of Escherichia coli. *Proc Natl Acad Sci U S A* **103**, 11754-11759
5. Khetrapal, V., Mehershahi, K., Rafee, S., Chen, S., Lim, C. L., and Chen, S. L. (2015) A set of powerful negative selection systems for unmodified Enterobacteriaceae. *Nucleic Acids Res* **43**, e83
6. Ryu, Y., and Schultz, P. G. (2006) Efficient incorporation of unnatural amino acids into proteins in Escherichia coli. *Nat Methods* **3**, 263-265
7. Murphy, K. C., and Campellone, K. G. (2003) Lambda Red-mediated recombinogenic engineering of enterohemorrhagic and enteropathogenic E. coli. *BMC Mol Biol* **4**, 11
8. Chang, A. C., and Cohen, S. N. (1978) Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**, 1141-1156
9. Weiss, D. S., Chen, J. C., Ghigo, J. M., Boyd, D., and Beckwith, J. (1999) Localization of FtsI (PBP3) to the septal ring requires its membrane anchor, the Z ring, FtsA, FtsQ, and FtsL. *J Bacteriol* **181**, 508-520
10. Ovchinnikov, S., Park, H., Varghese, N., Huang, P. S., Pavlopoulos, G. A., Kim, D. E., Kamisetty, H., Kyrpides, N. C., and Baker, D. (2017) Protein structure determination using metagenome sequence data. *Science* **355**, 294-298
11. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132
12. DeLano, W. L. (2002). PyMOL. DeLano Scientific, San Carlos, CA, 700.
13. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612
14. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-38, 27-38