Appendix

A novel pilus superfamily from the endospores of pathogenic Bacilli.

Brajabandhu Pradhan, Janine Liedtke, Mike Sleutel, Toril Lindbäck, Ephrem Debebe Zegeye, Kristin O'Sullivan, Ann-Katrin Llarena, Ola Brynildsrud, Marina Aspholm, Han Remaut

Table of contents:

	Ex vivo S-Ena (EMD- 11592)	*rec*ENA1B (EMD-11591) (PDB 7A02)
Data collection and processing	CryoARM300, BECM	CryoARM300, BECM
Magnification	60.000	60.000
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	62.5	64.66
Defocus range (μm)	-0.5 to -3.5	-0.5 to -3.5
Pixel size (Å)	0.82	0.784
Symmetry imposed	Helical Rise= 3.22937 Rotation=31.0338	Helical Rise= 3.43721 Rotation=32.3504
Initial particle images (no.)	53501	100495
Final particle images (no.)	42822	65466
Map resolution (Å)	3.2	3.05
FSC threshold	0.143	0.143
Map resolution range (Å)		3.05-3.65 [1]
Refinement		
Initial model used	NA	de novo
Model resolution (Å)	NA	2.81
FSC threshold	NA	0.143
Model resolution range (Å)		
Map sharpening *B* factor (Å^2)	25.9 B-iso of density modification	27.4 B-iso of density modification
Model composition		
Non-hydrogen atoms	NA	18699 [2]
Protein residues	NA	2576 [2]
Ligands	NA	0
B factors (Å^2)		
Protein	NA	54.39
Ligand	NA	NA
R.m.s. deviations		
Bond lengths (Å)	NA	0.008
Bond angles (°)	NA	0.736
Validation		
MolProbity score	NA	1.93
Clashscore	NA	8.07
Poor rotamers (%)	NA	0
Ramachandran plot		
Favored (%)	NA	101 (92%) [3]
Allowed (%)	NA	9 (8%) [3]
Disallowed (%)	NA	0 [3]

[1] Numbers reflect the density modified cryo-EM map calculated using ResolveCryoEM (Terwilliger et al., 2019)

[2] Numbers reflect a S-Ena model with 23 Ena1B protomers

[3] Numbers for a single Ena1B protomer

Appendix Table S2 Primers used in this work. To allow assembly of the PCR fragments, primers B and C contain sequences overlapping each other (italic).

Primer	Sequence (5'-3')	
Deletion mutants		**Background strain**
Δena1A		
A: 2184	AATGGCGCCAGTTCAATTAC	*B. cereus* NVH 0075-95
B: 2198	*CCTCTCTACATAGCCTTT*TCCCCTCTCTCTT	
C: 2199	*AAGGCTATGTAGAGAGGG*GAATTAGTAT	
D: 2178	CCTCCTATTCTCCCACCTGAAA	
Δena1B		
A: 2164	TCCATGTGGTATGGCAAAAA	*B. cereus* NVH 0075-95
B: 2165	*CCATATATTACA**T**ACTAATT*CCCCTCTC	
C: 2166	*AATTAGTATGTAATATATGG*TGATTTAAAGATT	
D: 2167	AACCTACTTGCCCCTGTCCT	
Δena1C		
A: 2200	CGCATCTTGTTTAGGTGCAA	*B. cereus* NVH 0075-95
B: 2201	*ATTTTTTTGTTATCCTTTTCA*TAAGACTGTTTAC	
C: 2202	*TGAAAAGGATAACAAAAAAAT*TATTGCTTTTG	
D: 2176	AGGTGGAGGGACAATCCAAAC	
Δena1AB		
A: 2164	TCCATGTGGTATGGCAAAAA	*B. cereus* NVH 0075-95
B: 2186	*CCATATATTACATAGCCTTTCC*CCTCTC	
C: 2197	*AAAGGCTATGTAATATATGG*TGATTTAAAGAT	
D: 2167	AACCTACTTGCCCCTGTCCT	
Δena1ABC		
A: 2164	TCCATGTGGTATGGCAAAAA	*B. cereus* Δena1AB
B: 2202	*TGAAAAGGATAACAAAAAAAT*TATTGCTTTTG	
C: 2201	*ATTTTTTTGTTATCCTTTTCA*TAAGACTGTTTAC	
D: 2200	CGCATCTTGTTTAGGTGCAA	
RT-PCR		Gene
2116/2117	AAGTGCGTCTAATCAACAAGGAAA/ GGGAAATCTCCCATGAACACA	*rpoB*
2176/2177	AGGTGGAGGGACAATCCAAAC/ GGCGAAACGTAAATGAAATGC	*ena1A*
2174/2175	CCACTGGAAGTAGCGCATCTT / GCCGCTGTTCCAAGAATTGT	*ena1B*
2178/1279	CCTCCTATTCTCCCACCTGAAA / CTCCAGCGAACTCATTGGTAACT	*ena1C*
2180/2181	GGGTGTACGAGGGTGATATGAATT/ TGTCGTTCCGCCAAGTGTT	*dedA*
Complementation		**Plasmid construct**
2220/2221	GCGGATGTTGTTGGACAA/ACGTGCAAACACATGAATCG	pHT315-*pena1AB*
2220/2222	GCGGATGTTGTTGGACAA/ TCCATTTGCACAACATGAAAG	pHT315-*pena1A*
2197/2185	AAAGGCTATGTAATATATGGTGATTTAAAGAT/CGCATCTTGTTTAGGTGCAA	pHT315-*pena1C*

Appendix Table S3 Constructs used and generated in this study.

Plasmids		
Name	**Details**	**Reference**
pena1A	pHT315::*ena1A* (complementation of *B. cereus Δena1A*)	*this study*
pena1AB	pHT315::*ena1AB* (complementation of *B. cereus Δena1B*)	*this study*
pena1C	pHT315::*ena1C* (complementation of *B. cereus Δena1C*)	*this study*
pMAD-I-SceI	Shuttle vector for making deletion mutant, carries I-SceI restriction site	(Lindback et al., 2012)
pMADΔena1A	pMAD-I-SceI carrying DNA sequences flanking *ena1A*	*this study*
pMADΔena1B	pMAD-I-SceI carrying DNA sequences flanking *ena1B*	*this study*
pMADΔena1C	pMAD-I-SceI carrying DNA sequences flanking *ena1C*	*this study*
pMADΔena1AB	pMAD-I-SceI carrying DNA sequences flanking *ena1AB*	*this study*
pMADΔena1ABC	pMAD-I-SceI carrying DNA sequences flanking *ena1ABC*	*this study*
pBKJ233	Express I-SceI enzyme	(Janes and Stibitz, 2006
pMS-BcE1A	pET-28a(+) with CDS *rec*Ena1A	*this study*
pMS-BcE1B	pET-28a(+) with CDS *rec*Ena1B	*this study*
pMS-BcE1C	pET-28a(+) with CDS *rec*Ena1C	*this study*
Proteins & coding sequences		
***rec*Ena1A**		

Coding sequence	ATGCATCACCATCACCATCACAGCAGCGGTGAAAATCTGTATTTTCAGGGCGCTTGTGAATGTAG CAGCACAGTCCTGACCTGTTGTTCGGACAACTCTAGTAATTTTGTGCAGGATAAAGTTTGCAACC CCTGGTCTTCCGCGGAAGCAAGCACTTTCACCGTATATGCGAACAATGTTAACCAAACATTGTC GGCACCGGCTATTTAACATATGATGTGGGGCCCGGTGTGAGCCCGGCCAACCAGATTACGGTTAC GGTTCTGGACTCCGGCGGCGGTACTATCCAGACCTTTCTGGTCAACGAAGGGACGTCTATCTCAT TTACCTTTCGCCGCTTTAACATTATTCAGATTACAACCCCAGCCACACCGATTGGCACGTATCAG GGTGAATTTTGCATCACGACCCGGTATTTAATGGCCTAA

Protein sequence	MHHHHHHSSGENLYFQGACECSSTVLTCCSDNSSNFVQDKVCNPWSSAEASTFTVYANNVNQNIVGTGYLT YDVGPGVSPANQITVTVLDSGGGTIQTFLVNEGTSISFTFRRFNIIQITTPATPIGTYQGEFCITTRYLMA

***rec*Ena1B**	

Coding sequence	ATGCACCACCACCACCATCATTCTAGCGGTGAAAACCTGTACTTTCAGGGTAACTGCAGCACCAA TCTGTCATGCTGTGCCAATGGTCAGAAGACCATTGTCCAGGATAAAGTCTGCATCGACTGGACCG CAGCCGCTACTGCAGCAATCATTTACGCTGATAATATCAGCCAAGACATCTACGCTTCAGGCTAT CTGAAAGTGGATACAGGTACGGGTCCCGTGACCATCGTCTTTTACTCTGGTGGAGTCACAGGCAC CGCTGTGGAGACCATTGTGGTCGCCACGGGTTCGTCGGCCAGCTTTACGGTGCGCCGTTTTGATA CCGTCACTATTCTGGGCACCGCAGCAGCGGAGACTGGTGAGTTTTGTATGACCATCCGTTACACT TTGAGCTAA

Protein sequence	MHHHHHHSSGENLYFQGNCSTNLSCCANGQKTIVQDKVCIDWTAAATAAIIYADNISQDIYASGYLKV DTGTGPVTIVFYSGGVTGTAVETIVVATGSSASFTVRRFDTVTILGTAAAETGEFCMTIRYTLS

***rec*Ena1C**	

Coding sequence	ATGCATCATCATCATCACCATTCCAGCGGTGAGAATCTCTATTTCCAAGGCAAACCGCACAAAAA TATCGGCTGCTTTGCGCCGCTCTCCATTATCTGCCAGCCGACCTGTCCGTGCCCGCCGCCAATTT TACCGCCGGAACGCGGTGACGCCGAGCTGGTCACAAATGAATTTGCGGGGGACATCCTGATTAGC AACGATTTTATTCCAATTAGCCAGAAACAGCTGAAACAAACAAACACCACCGTTAATATCTGGAA AAACGACGGAATCGTTTCCCTGAGCGGCACGATTTCAATTTATAATAATCGCAATTCGACCAACG CGCTGTCGATTCAGATTATCAGTAGTACGACCAATACCTTTACAGCGCTCCCGGGGAATACGATT TCCTATACTGGTTTTGACCTGCAGTCCGTCTCTGTTATCGACATTCCAAGCGATCCAAGTATCTA CATTGAGGGCCGCTATTGTTTTCAGTTAACTTACTGTAAATCTAAACGCGATTGTCTTTAA
Protein sequence	MHHHHHHSSGENLYFQGKPHKNIGCFAPLSIICQPTCPCPPPILPPERGDAELVTNEFAGDILISNDFIPISQKQL KQTNTTVNIWKNDGIVSLSGTISIYNNRNSTNALSIQIISSTTNTFTALPGNTISYTGFDLQSVSVIDIPSDPSIYIEG RYCFQLTYCKSKRDCL

Appendix Table S4 Overview of species included in the pairwise tBLASTn searches for the individual query proteins.

Species	Number of genomes
B. albus	1
B. anthracis	63
B. bombysepticus	1
B. cereus	85
B. cytotoxicus	14
B. manliponensis	1
B. gaemokensis	2
B. luti	1
B. mycoides	33
B. mobilis	5
B. nitratireducens	6
B. pacificus	3
B. paramycoides	2
B. paranthracis	3
B. pseudomycoides	8
B. proteolyticus	1
B. subtilis	127
B. thuringiensis	50
B. toyonensis	204
B. tropicus	7
B. wiedmannii	119
Total	**735**

Appendix Table S5 Sequence and origin of query proteins used to search for homo- and orthologs within subject genomes in Table S4 using parwise tBLASTn searches.
EnaX_NM_Oslo: Query protein sequence is sequenced from amplified lab strain.
EnaX_GCF_001044825: Query sequence protein is from the gene in the publicly available NVH 0095-75:
GCF_001044825.1_ASM104482v1. NM 0095-75

PCR product	
Ena1B_NM_Oslo	MGNCSTNLSCCANGQIIVQDKVCIDWTAAATAAIIYADNISQDIYASGYLKVDTGTGPVTIVFYSGGVT GTAVETIVVATGSSASFTVRRFDTVTILGTAAAETGEFCMTIRYTLS
From the GCF_001044825.1:	
Ena1A_GCF_0010 44825, KMP91697.1	MACECSSTVLTCCSDNSSNFVQDKVCNPWSSAEASTFTVYANNVNQNIVGTGYLTYDVGPGVSPANQ ITVTVLDSGGGTIQTFLVNEGTSISFTFRRFNIIQITTPATPIGTYQGEFCITTRYLMA
Ena1C_GCF_0010 44825, KMP91699.1	LKPHKNIGCFAPLSIICQPTCPCPPPILPPERGDAELVTNEFAGDILISNDFIPISQKQLKQTNTTVNIWKN DGIVSLSGTISIYNNRNSTNALSIQIISSTTNTFTALPGNTISYTGFDLQSVSVIDIPSDPSIYIEGRYCFQLTY CKSKRDCL

References

Argimon, S., Abudahab, K., Goater, R.J.E., Fedosejev, A., Bhai, J., Glasner, C., Feil, E.J., Holden, M.T.G., Yeats, C.A., Grundmann, H*., et al.* (2016). Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. Microb Genom *2*, e000093. Burnley, T., Palmer, C.M., and Winn, M. (2017). Recent developments in the CCP-EM software suite. Acta Crystallogr D Struct Biol *73*, 469-477.

Farabella, I., Vasishtan, D., Joseph, A.P., Pandurangan, A.P., Sahota, H., and Topf, M. (2015). TEMPy: a Python library for assessment of three-dimensional electron microscopy density fits. J Appl Crystallogr *48*, 1314-1323.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximumlikelihood trees for large alignments. PLoS One *5*, e9490.

Terwilliger, T.C., Ludtke, S.J., Read, R.J., Adams, P.D., and Afonine, P.V. (2019). Improvement of cryo-EM maps by density modification. bioRxiv.