

Acta Crystallographica Section D

Volume 70 (2014)

Supporting information for article:

The type IV secretion protein TraK from the *Enterococcus* conjugative plasmid pIP501 exhibits a novel fold

Nikolaus Goessweiner-Mohr, Christian Fercher, Karsten Arends, Ruth Birner-Gruenberger, Diana Laverde-Gomez, Johannes Huebner, Elisabeth Grohmann and Walter Keller

Table S1 3D alignments with DALI (A) and MATRAS (B).

Entry number	PDB entry code	z-score	RMSD (A) /	Aligned amino	Description
A – 1	3kzt	6.1	3.5	109	Uncharacterized protein
A – 2	3rob	4.8	3.6	95	Uncharacterized conserved protein
A – 3	43c6	4.6	3.6	92	Putative uncharacterized protein
A – 4	3cnx	4.5	3.8	90	Uncharacterized protein
A – 5	1jkg	4.4	4.1	100	P15
A – 6	3fka	4.4	3.7	90	Uncharacterized NTF-2 like protein
A – 7	1jn5	4.4	4.0	99	P15
A – 8	2f86	4.4	3.8	92	Hypothetical protein K11E8.1D
A – 9	2w2c	4.4	4.5	95	Calcium/Calmodulin-dependent protein kinase
A – 10	2ux0	4.4	4.5	95	Calcium/Calmodulin-dependent protein kinase
B – 1	3kzt	18.08	16.02	40-166	Uncharacterized protein
B – 2	3a76	17.89	13.0	1-133	Gamma-hexachlorocyclohexane
B – 3	3mwx	13.66	2.9	89-165	Aldose 1-epimerase
B – 4	3lod	10.47	5.7	-1-81	Putative acyl-CoA N-acyl-transferase
B – 5	3os7	10.11	2.4	96-172	Galactose mutarotase-like protein
B – 6	3l4y	9.32	0.5	40-106	Maltase-glucoamylase, intestinal
B – 7	2r7h	8.96	5.0	18-102	Putative D-alanine N-acetyl-transferase of
B – 8	3jvn	8.45	5.5	3-97	Acetyl-transferase
B – 9	3db7	8.04	4.2	0-76	Putative calcium-regulated periplasmic protein
B – 10	1n71	8.02	4.2	1-79	AAC(6')-II

The first 10 alignment entries are listed according to their respective z-score. The PDB entry code, z-score, RMSD (A) or Rdis (B), aligned amino acids and a description is given for each entry.

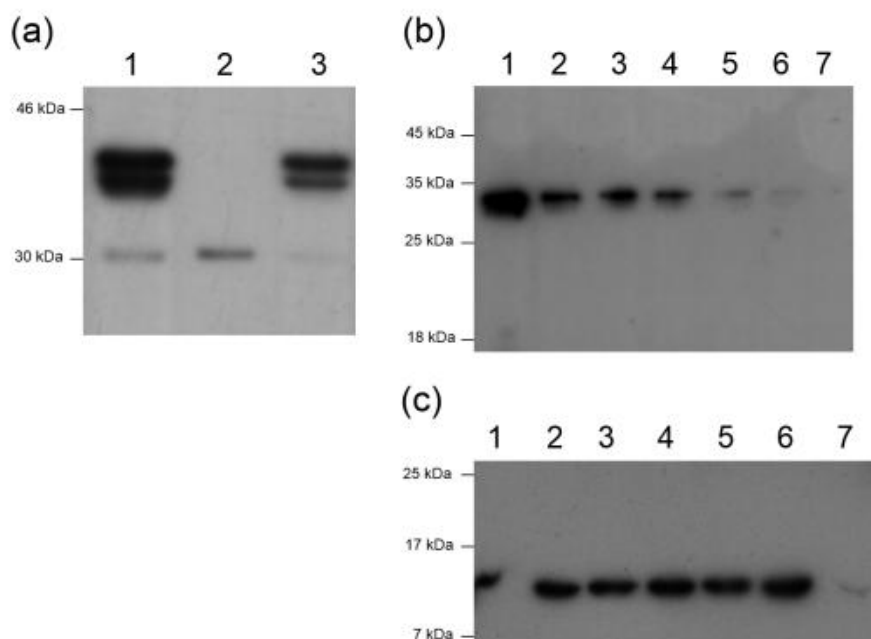


Figure S1 Analysis of expression and localization of TraK. (a) *traK* start codon mutagenesis: Western blot analysis of *E. coli* BL21-CodonPlus (DE3)-RIL pQTEV-*traK* cells expressing TraK. The detected triplet corresponds to the three different gene products originating from the start codon provided by pQTEV and two alternative *traK* start codons, respectively (lane 1 - 2 μ l cell lysate, lane 3 - 1 μ l cell lysate). Mutagenesis of the pQTEV-derived and the first *traK* start site clearly demonstrated utilization of the second ATG within the *traK* *orf* (lane 2). (b and c) Protease protection assay with anti TraK and anti TraN polyclonal antibodies: Protease digestion demonstrated that the C-terminal domain of TraK is extracellular. *E. faecalis* JH2-2 protoplasts expressing the pIP501 *traK* genes were subjected to a protease protection assay. Beside TraK (b), the cytoplasmic transfer protein TraN (c) was analyzed in the same sample as a control. Protoplasts were either intact (lanes 1 to 6) or fully permeabilized with Triton X-100 (lane 7). Lanes represent different concentrations of proteinase K: lane 1 - no proteinase K; lane 2 - 0.1 μ g ml^{-1} ; lane 3 - 0.4 μ g ml^{-1} ; lane 4 - 1 μ g ml^{-1} ; lane 5 - 5 μ g ml^{-1} ; lane 6 - 10 μ g ml^{-1} ; lane 7 - 5 μ g ml^{-1} + 1% Triton X-100. Molecular weight marker (a) and (c): Prestained Protein Marker, Broad Range (NEB, Ipswich, MA, USA); Molecular weight marker (b): Unstained Protein Molecular weight marker (ThermoScientific, Waltham, Massachusetts, USA).

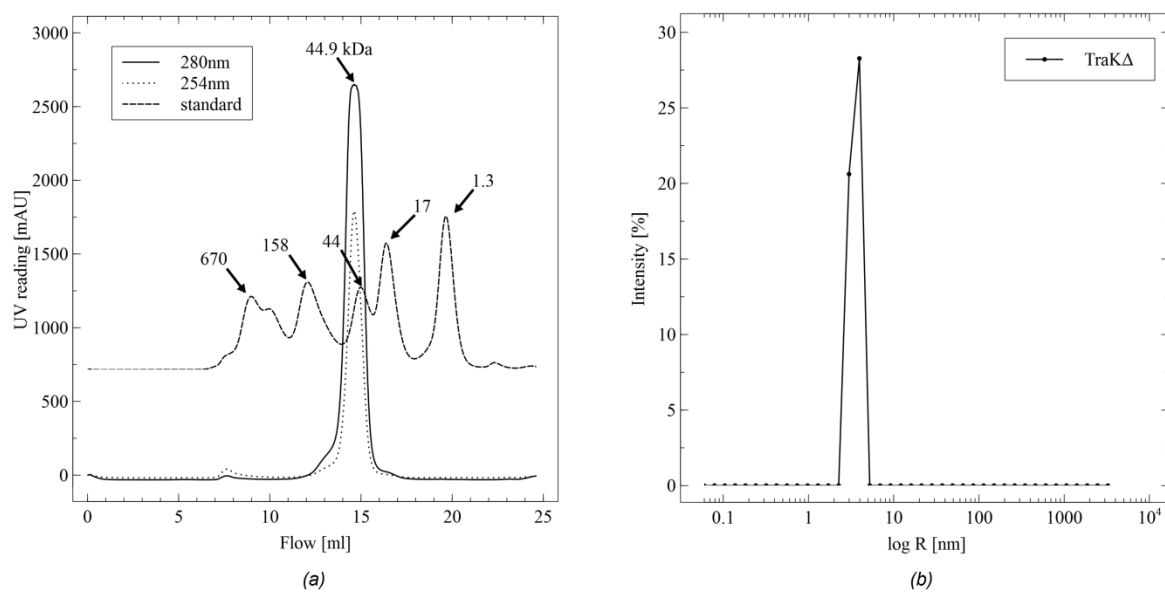


Figure S2 Biophysical characterization of TraKΔ. (a) TraKΔ elutes as a single peak from the Superdex 200 size exclusion column. The 280 nm (black) and 254 nm (dotted) readings are shown. A standard (BioRad) is shown with its molecular weight (discontinued). (b) In the monodispersity analysis via DLS TraKΔ appears as a narrow peak.

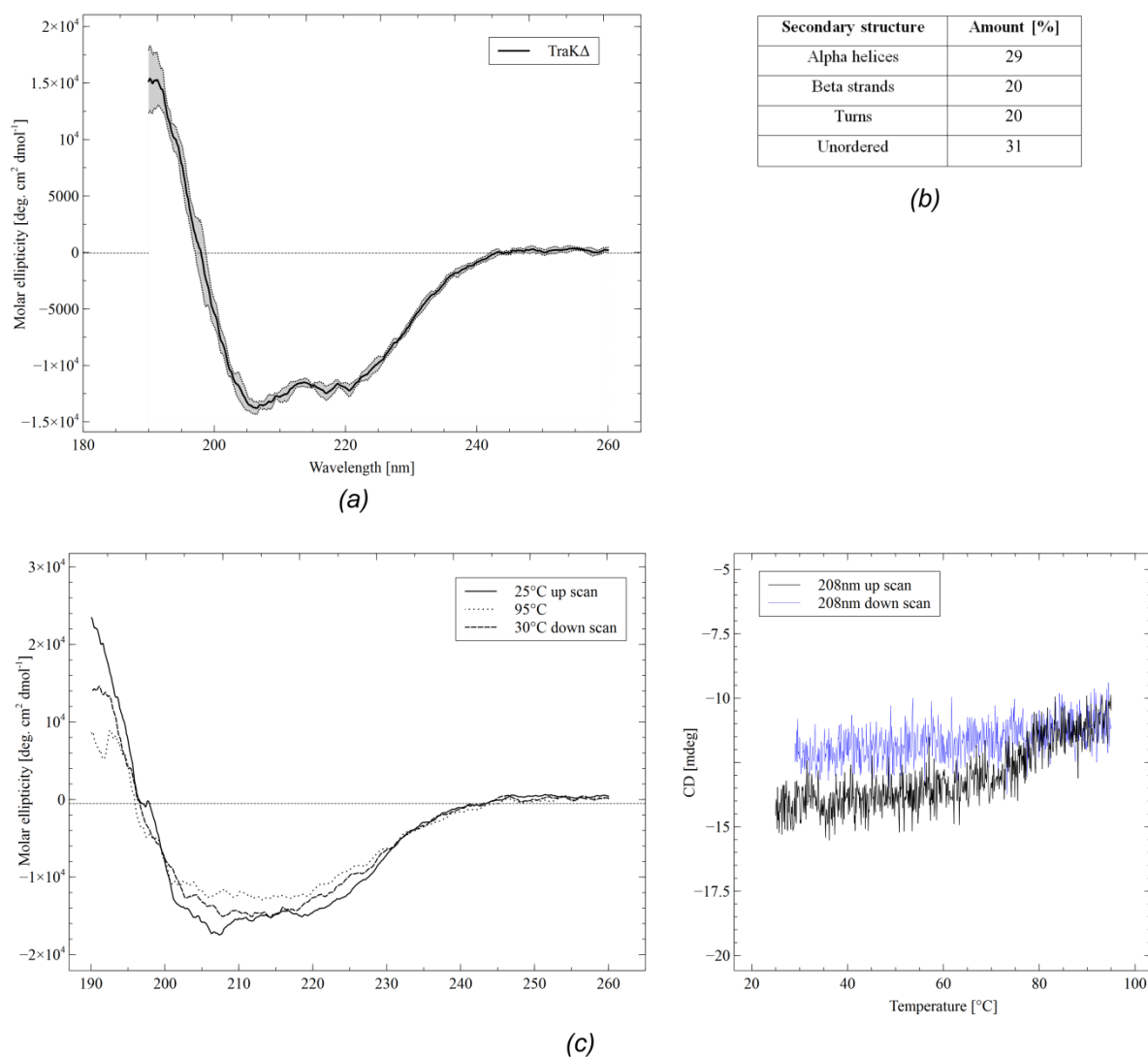


Figure S3 CD analysis of TraKΔ. (a) TraKΔ is folded in solution. The black curve represents the average of 10 individual wavelength scans. The standard deviation is displayed as grey area. (b) Secondary structure content of TraKΔ. The NRMSD is 0.022. (c) TraKΔ unfolding and refolding characteristics. The CD spectra are shown at 25 and 95 °C and after cooling to 25 °C (left panel). The temperature scan (up- and down-scan) is shown in the right panel.

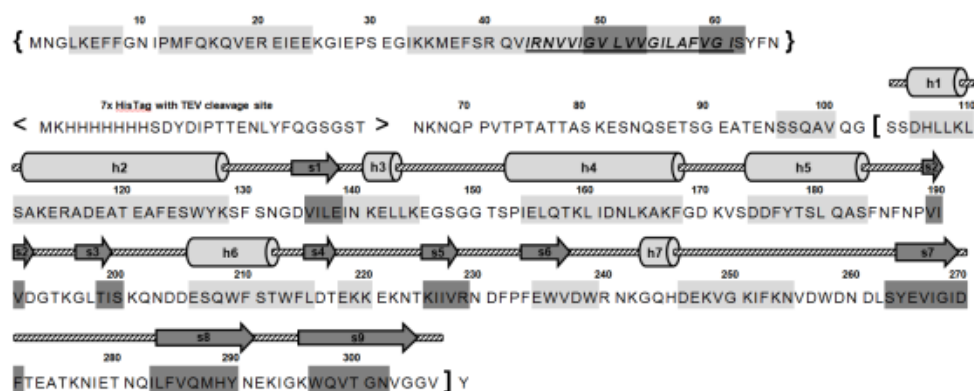


Figure S4 Overview of the full length TraK amino acid sequence and the TraK Δ construct.

Secondary structure features predicted by PsiPred are marked in light grey (helices) and dark grey (sheets); the N-terminal domain of the full length protein, not present in the TraK Δ construct, is given in curved brackets; italic + bold and underlined amino acids mark the predicted trans-membrane region of TraK; according to MS measurements, the full length TraK Δ construct is present in the TraK Δ crystals; square brackets highlight the sequence present in the actual X-ray structure; secondary structure elements found in the X-ray model are shown as light grey tubes (helices) and dark grey arrows (sheets).



Figure S5 Sequence alignment of TraK-like T4SS proteins. The residues are highlighted according to their conservation based on the Blosum26 score. All of the analyzed proteins showing a TraK-like secondary structure composition belong to T4SSs of G⁺ origin (five plasmids, three ICEs). Except of one protein from *S. pyogenes* plasmid pSM19035, all candidate T4SS proteins were found in *Enterococcus* species.

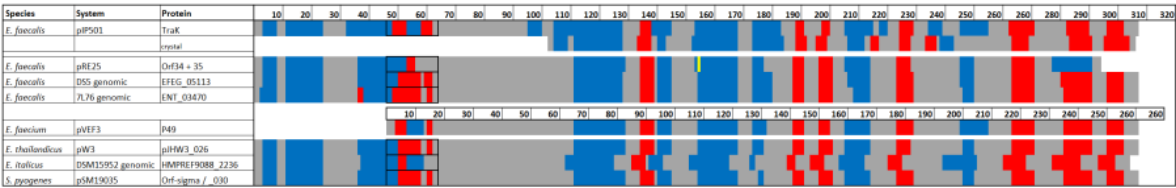


Figure S6 Secondary structure based alignment of TraK-like proteins. Secondary structure (Pspred) and trans-membrane motif (HMMTOP) prediction for TraK-like proteins of plasmids and ICEs; alpha helices (blue), beta strands (red) and trans-membrane motifs (boxes) are highlighted; a yellow box marks where pRE25 proteins ORF34 and ORF35 were joined.