

Structural and biochemical properties of *Francisella tularensis* lipoprotein FTT1103

¹L. Hernychova, ²E. Chovancova, ²J. Damborsky, ^{1,3}L. Balonova, ¹I. Pavkova, ⁴M. Simsova, ^{3,4}M. Capkova, ³Z. Bilkova, ⁵M. Novotny, ¹A. Macela, ¹J. Stulik

¹Faculty of Military Health Sciences, University of Defence, Hradec Kralove, Czech Republic
²Loschmidt Laboratories, Faculty of Science, Masaryk University, Brno, Czech Republic
³University of Pardubice, Pardubice, Czech Republic
⁴Apronex s.r.o., Jesenice u Prahy, Czech Republic
⁵Department of Chemistry, Indiana University, IN, USA

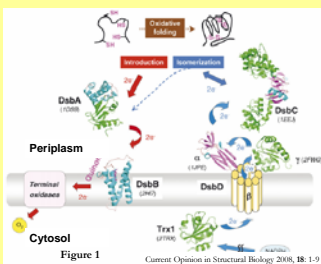


Introduction

Francisella tularensis and lipoprotein FTT1103

Francisella tularensis is a highly virulent Gram-negative bacterium that causes tularemia, a disease that is often fatal if untreated. Little is known about the components of the microorganism that are responsible for inducing disease in its hosts. Based on the recent genomic and proteomic studies was found the conserved hypothetical lipoprotein FTT1103 with virulence and vaccine potential. This protein exhibits high homology with proteins of the disulfide oxidoreductase DsbA family catalyzing disulfide bond formation. The DsbA protein is essential for proper folding of many extracytosolic proteins including toxins and other virulence factors in other pathogens.

Figure 1: The oxidation and reduction/isomerization pathways of oxidative protein folding in *E. coli*. The enzymes DsbA and DsbC interact with redox regulators DsbB and DsbD.



Structural analysis

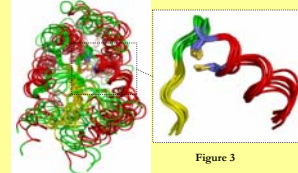
Bioinformatics analysis

Primary and secondary structure analysis

InterProScan and Motif-Scan servers and Hmmpfam and HHSearchCDD tools accessible via the GeneSilico meta-server were used to identify putative patterns and domains within the query proteins.

Results of secondary database searches indicate that protein FTT1103 probably contains two different domains and a signal peptide located on its N-terminus (Figure 1). The signal peptide consists of the residues 1-21 and was identified as a part of PROSITE profile described as prokaryotic membrane lipoprotein lipid attachment site (PROSITE accession: PS51257). The signal peptide is followed by Cys22 predicted as a potential binding site for a lipid group (palmitate or diacylglyceride). The N-terminal domain of FTT1103 may have oligomerization and/or chaperone activity and C-terminal domain is essential for the catalytic function of DsbA proteins.

Figure 2: Two different domains were predicted within the sequence of FTT1103: 1) Domain amino terminal to FKBP-type peptidyl-prolyl isomerase (FKBP_N, Pfam accession: pfam01346, CDD accession: COG0545) localized in the N-terminal part of the sequence with the start around positions 25-35 and 2) DSB-like thioredoxin domain (Pfam accession: pfam1323) comprising the C-terminal part of the sequence with the start around positions 135-145.



Both conservation analysis and analysis of 3D structures related to FTT1103 were conducted to identify highly conserved residues with potential structural or functional importance. Highly conserved residues included both cysteines essential for the catalytic function of DsbA proteins (Cys164 and Cys167 of FTT1103).

Figure 3: Structural superposition of the proteins related to FTT1103 including PDB-IDs 3be1, 1v57, 1z6m, 1t5b, 1ecj, 2hi7 chain A, 1bed and 3c7m. Loops are in green, α -helices are in red and β -strands are in yellow. Functionally important cysteines, corresponding to Cys164 and Cys167 of FTT1103, are highlighted as blue-yellow "sticks".

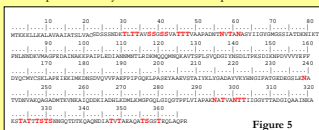
Localization

FTT1103 was annotated according to UniProt database as **outer membrane protein bound in the periplasmic space**. Moreover another programs (SignalP, SOSUI and Psortb) support its membrane localization. Although the proteomic studies indicate that FTT1103 can be a membrane protein, its localization in the outer membrane has not been experimentally verified. Our experiments based either on the separation of outer and inner leaflet of bacterial membrane in sucrose density gradient or production of the recombinant FTT1103-PhoA fusion proteins confirmed its localization in the cytoplasmic membrane and the periplasmic space.

Figure 4: Design of supposed localization of FTT1103 lipoprotein. A lipid group is located in N-terminal part of the sequence (attached via Cys22) and it might serve for anchoring of FTT1103 to the membrane. The C-terminal part of the protein should be localized in the periplasmic space.

Putative glycosylated sites

The FTT1103 sequence was searched for the presence of N- and O-glycosylation sites using six different prediction methods (PROSITE database, EnsembleGly, GPP, NetOGlyc and NetNGlyc servers). Potential glycosylation sites (Asp, Ser and Thr residues) were ordered according to the number of methods predicting them as glycosylated and the score from the ensemble of support vector machine methods. It is important to note that all currently available methods were developed for identification of glycosylation sites in mammalian proteins and use the rules that might be not directly applicable to prokaryotic systems. The results should be therefore interpreted with caution. The sites predicted as glycosylated by at least two methods in FTT1103 are summarized in Figure 5 (potential glycosylated sites are marked as red letters). Two asparagines - Asn239 and Asn55 - were predicted as glycosylated by all six methods and three more asparagines - Asn300, Asn295 and Asn59 - by five methods. Thr34, Thr45, Thr46 and Ser37 predicted by four methods represent the best candidates for O-glycosylation sites.



Relative solvent accessibility (RSA) of FTT1103 amino acid residues was predicted by SABLE, Jnet and ACCpro methods accessible via the GeneSilico meta-server and the individual residues were classified as buried or exposed. Three different RSA thresholds of 25%, 5% and 0% were used for the Jnet and ACCpro classifications. All putative glycosylation sites were classified as **surface exposed**.

Summary

This study confirmed:

- 1) Presence of DSB-like thioredoxin domain and its catalytic activity
- 2) Localization of protein in periplasmic space
- 3) Presence of glycosylation and acylation

Our study provides insights to the structure and function of protein FTT1103, which can serve for development of efficient vaccine.

Catalytic activity

Catalytic function

with similarity to the disulfide oxidoreductase DsbA family and N-terminal domain of FKBP-type peptidyl-prolyl isomerase, respectively. Catalytic function was predicted and experimentally confirmed for Cys164 and Cys167, which are directly involved in the reaction mechanism of DsbA.

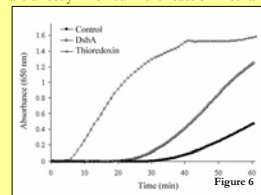


Figure 6: Disulfide oxidoreductase activity of the *F. tularensis* DsbA homolog determined by the insulin reduction assay in the absence and presence of recombinant *E. coli* thioredoxin or recombinant *F. tularensis* DsbA homolog. The background reduction by DTT without any enzyme present was used as a control.

Method: Reaction mixture: 150 μ M insulin in 0.1M potassium phosphate buffer, pH 7.0 + 2 mM EDTA + 5 μ M thioredoxin *E. coli* or 5 μ M rDsbA *F. tularensis* (Apronex) + 0.33mM DTT. Absorbance at 650 nm was measured every 30s for 1 hour

Localization in *E. coli*

Topology study

The genetic engineering methods were used for validation of topology data obtained in the bioinformatic study. A set of vectors for expression of fusion recombinant proteins as well as their free variants was prepared for this reason (Figure 7). Expression vectors were prepared by "in frame" cloning of FTT1103 gene and reporter gene for periplasmatically active enzyme alkaline phosphatase (PhoA). Activity of PhoA due to enzyme secretion into periplasmic space in the presence of functionally active N-terminal signal sequence was used to determine lipoprotein localization in *E. coli*.

- 1 - FTT1103-AD-PhoA: fusion protein FTT1103 with signal peptide and with alkaline phosphatase
- 2 - FTT1103-PhoA: fusion protein FTT1103 without signal peptide and with alkaline phosphatase
- 3 - PhoA: alkaline phosphatase
- 4 - FTT1103-AD: protein FTT1103 with signal peptide

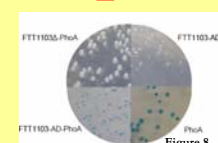


Figure 7: A set of recombinant proteins:

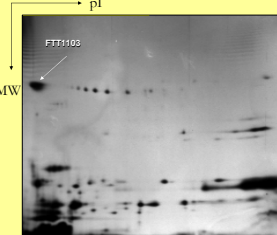
Expression of alkaline phosphatase activity was assessed on LB agar plates containing 5-bromo-4-chloro-3-indolylphosphate as PhoA substrate. **Coloured colonies** were observed in the case of **PhoA localization in periplasmic space** - alkaline phosphatase (PhoA) and fusion protein FTT1103 with signal peptide and with alkaline phosphatase (FTT1103-AD-PhoA), **colourless colonies** were detected in the case of **PhoA localization in cytoplasm** - protein FTT1103 with signal peptide (FTT1103-AD) and fusion protein FTT1103 without signal peptide and with alkaline phosphatase (FTT1103-PhoA), see **Figure 8**.

Posttranslational modification

Acylation

According to prediction programs the FTT1103 gene encodes putative lipoprotein and in addition Cys22 was predicted as a putative lipid attachment site. Presence of acylation was experimentally verified by ³H palmitate labeling during cultivation of bacteria in order to label lipid moieties of newly synthesized bacterial lipoproteins. Membrane-associated proteins isolated from the labeled bacteria were then separated by 2-DE and exposed to autoradiographical film.

Figure 9: Representative 2D autoradiographical pattern of [³H] palmitic acid labeled lipoproteins presented in *F. tularensis* membrane fraction.



Glycosylation

Bioinformatics analysis of the FTT1103 protein predicted five potential sites for N-glycosylation and four for O-glycosylation. Protein glycosylation was confirmed by the mass and charge shift in the protein species on 2-DE protein pattern of the deletion mutant for gene glycosyl transferase *pilO2*. This gene codes enzyme that catalyze the polymerization of the carbohydrate chains of peptidoglycan. Protein spots of the FTT1103 were shifted in both pI and Mw (Figure 10). This proteins could be modified by glycosylation.

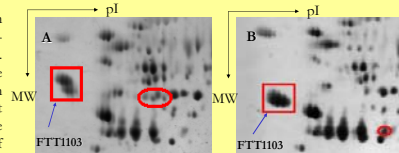
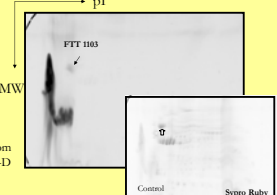


Figure 10: Proteomic comparison between parental *F. tularensis* FTS200 and *pilO2* mutant strains using two-dimensional gel electrophoresis. Red lines show qualitative protein changes. Enlarged regions of representative silver stained 2D maps of proteins from membrane enriched fractions A) FTS200 and B) *pilO2*.

The detection and identification of putative glycoprotein FTT1103 was further verified by using the Pro-Q Emerald 300 glycoprotein stain (Figure 11) and lectin affinity chromatography.

Figure 11: Mini 2-D gel (pH 3-10) of membrane proteins isolated from *F. tularensis* strain Schu4 with detected glycoprotein and the identical mini 2-D gel stained with SYPRO Ruby (Control).



Future pains

- 1) Finding of active sites in the second FKBP-type peptidyl-prolyl isomerase domain which are necessary for function of protein
- 2) Electron microscopy study for confirmation of localization N-part of protein in the outer bacterial membrane
- 3) Glycomic profiling by chemical release of glycans

Acknowledgement: This work was supported by grants from Ministry of Education ME08105, Ministry of Defence FVZ000604 and FVZ200802, and Czech Science Foundation GA203/09/857.