Final report K-125275

Structure – function relationship of endotoxins

(Endotoxinok szerkezet – funkció kapcsolata)

Bacterial endotoxins are surface lipopolysaccharide (LPS) macromolecules comprising three defined regions distinguished by their genetics, structures, and function. These are the "lipid A", the "core oligosaccharide" and the "O-chain" polysaccharide portion. Some endotoxins lack the polysaccharide portion, these are called lipooligosacchirides (LOS). The hydrophilic carbohydrate part of LPSs is responsible for O-specific immunogenicity, while the hydrophobic lipid A anchor (composed of a phosphorylated diglucosamine backbone carrying varying numbers of fatty acids) is associated with endotoxicity. Once released from dead bacteria, endotoxins play a key role in the innate immune response during clinical infection with Gram-negative microbes. In mammalian hosts, they activate both the innate and the adaptive immune system, which can lead to septic shock and even death, whereas in plants they activate the innate immunity defense system, also called basal defense. The analysis of native lipid A structures and structural modifications in relation to biological activity assays and environmental stresses, as well as the better understanding of the genetic background of the biosynthesis of LPS to find candidate targets for new antimicrobial agents are current emerging issues within the endotoxin research.

We have established two new setups for the quantitation, separation, and simultaneous characterization of bacterial lipid A components – including isobars – contained in highly heterogenic lipid A preparations from *E. coli, Salmonella, and Proteus bacteria*. One setup is based on an on-line reversed-phase-HPLC-MS/MS method and the other on a unique on-line non-aqueous capillary electrophoresis (NACE)-MS/MS methodology. The interface design, separation selectivity, and detection sensitivity differ greatly for the two techniques. While the RP-HPLC method is a valuable analytical tool for the profiling of complex lipid A samples based on differences in the lipid A's hydrophobicity (i.e., the number of fatty acids attached to the diglucosamine backbone), the NACE-MS/MS method is capable to separate and identify phosphorylation positional isomers of bacterial lipid A (and also some acylation isomers). With this latter technique, the separation and full structural identification of C1- and C4'-monophosphorylated lipid A isomers, having the same *m/z* and molecular formula – but different biological activities –, in bacterial lipid A mixtures becomes possible. Moreover, both normal and reverse CE polarities can be applied, which provide an increased sensitivity for the detection of minor lipid A components.

We have thoroughly investigated the tandem mass spectrometric profiles obtained by low-energy collision induced dissociation of four synthetic lipid A analog compounds and native lipid A from *E. coli* O83 bacterium with three different mass spectrometers (ESI-QTOF, ESI-QqQ, ESI-ion trap) to explore fragmentation pathways under low-energy CID conditions in the negative ion mode but - as a novelty - in the positive-ion mode, as well. We have proposed detailed fragmentation mechanisms for all type of primary and secondary fatty acid cleavages, as well as sugar ring fragmentations. New fragmentation rules have been discovered, by which C1- and C4'-phosphorylation isomers, as well as C3- and C3'-primary ester-linked acylation isomers of lipid A can be distinguished in a single sample. The position of a single phosphate

group of monophosphorylated lipid A molecules can be directly and unambiguously determined by MS/MS measurements in positive ion mode.

We started to analyze purified lipooligosaccharide extracts obtained from *Salmonella* and *Shigella* bacteria with the newly developed RP-HPLC–MS/MS method. Based on our preliminary results, the type and location of some non-stoichiometric, but biologically important substituents of the separated LOS components can be identified.

Pseudomonas aeruginosa is a priority target pathogen for antibiotic research and development because of its high resistance to a wide range of antibiotics. Acquisition of bacterial resistance is often associated with modifications of the lipid A component present in the Gram-negative outer membrane. For a deeper understanding of subtle chemical variations of lipid A in relation to biological properties, we performed energy-dependent mass spectrometric (ERMS) fragmentation in both negative and positive ionization modes, using our newly developed NACE-ESI-QTOF MS/MS technique for the identification of different lipid A isomeric forms present in the *Pseudomonas aeruginosa* bacterium. Several isobars could be identified that have not been detected earlier with HPLC-MS/MS by others. The use of the complementary positive and negative ion modes enabled the unequivocal assignment of the phosphorylation site and the position of acyl chains in lipid A compounds of three acylation families ranging from tetra- to hexa-acylation. Moreover, C1-monophosphorylated lipid A species were identified in *P. aeruginosa* for the first time.

The temperature dependence of the LPS structures obtained from *Pseudomonas* species was also investigated with three techniques: direct injection MS, HPLC-MS/MS, and the optimized NACE-MS/MS method. Separation of both phosphorylation and acylation isomers was most efficient by mass spectrometry coupled with capillary electrophoresis. The lipid A pattern showed quantitative differences depending on the growth temperature of the bacterium at 25 °C, 37 °C and 42 °C. At lower culture temperatures, the bacterium produces higher amounts of the more toxic, hexa-acylated components. Furthermore, components with more hydroxyl groups are present in smaller amounts at higher temperatures.

Our new micro-isolation technique (only 1 day long, and using only 10 cononies from agar plate) together with MALDI-MS measurements based on heat-inactivation of bacteria has been tested on endotoxins from three model bacteria: *Salmonella minnesota* and *Shigella sonnei*, with well-known structures of their lipooligosaccharides (LOSs), and lipid A part from *Escherichia coli*. The cells were grown either on agar plate or in liquid broth culture. The MALDI-TOF mass spectra of the such crude cell suspensions closely resembled to the spectra of the purified LOS and lipid A analogs. The fast and easy micromethod is suitable for extracting small quantities of LOS and lipid A directly from bacteria grown in small amount of culture media, and evaluate the structures with mass spectrometry. We have also tested a commercially available micro-isolation kit (LPS Extraction Kit from iNtRON Biotechnology DR) for the isolation of *Shigella sonnei* LOSs. However, it resulted in low yield of LOS and loss of phosphoetanolamine and 4-amino-4-deoxy-L-arabinose from the structures. Thus, we conclude that this kit cannot be used for LPS structure–function analysis, as loss of these biologically important substituents hinders the determination of active two-component system.

The lipid A chemical composition, endotoxin activity and inflammatory potential of subgingival biofilm LPS extracts from patients with periodontal diseases and healthy persons have been examined and compared with the HPLC-MS/MS method. Also, the endotoxin lipid A activities in subgingival plaque from patients with periodontal disease before and 3 months

after non-surgical periodontal treatment and healthy controls were investigated. The mass spectral profiles of the three samples revealed structural differences in the lipid A (mainly considering the phosphorylation pattern). This is the first study to consider the structure-function relationship of different lipid A isoforms from *Porphyromonas gingivalis* bacterium present in the subgingival niche. We propose that the lipid A activity of the plaque samples obtained from healthy and periodontitis patients could be a diagnostic and monitoring marker for maintenance of the periodontal condition.

We performed genetic analysis of two *Shigella sonnei* strains belonging to the same mutant line (*S. sonnei* 4303 and *S. sonnei* 4351). *S. sonnei* 4351 was generated by random mutagenesis from *S. sonnei* 4303 and selected based on its special LOS structure (Figure 1).

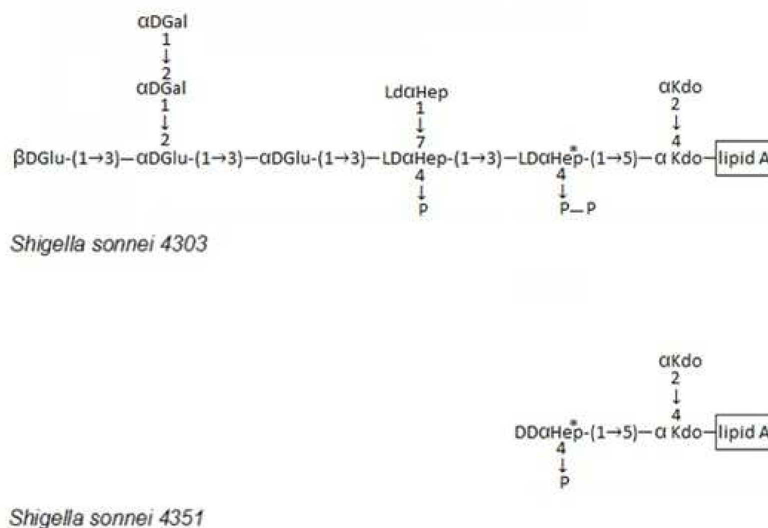

Figure 1. Lipooligosaccharide structures of *S. sonnei* 4303 and *S. sonnei* 4351.

LOS structure of *S. sonnei* 4351 is significantly shorter and it contains D-glycero-D-manno-heptose (D,D-heptose) instead of L-glycero-D-manno-heptose (L,D-heptose) in the core part. Phenotypic analysis of the strains showed differences between the optimal niche of *S. sonnei* 4303 and *S. sonnei* 4351. *S. sonnei* 4351 is more sensitive to Erythromycin, and Cefalexin, and its optimal range of growth temperature is also narrower.

Whole genome sequencing with IonTorrent PGM revealed the 4,5 Mbp long genomes (Table 1) with similar structures.

Descriptive statistics	*Shigella sonnei* 4303	*Shigella sonnei* 4351
Genome size (bp)	4.530.501	4.572.961
Hypothetic protein encoding sections (nr)	4554	4707
Number of rRNA genes	10	10
Number of tRNA genes	60	68
CRISPR region	1	1

Table 1. General statistics obtained for *Shigella sonnei* 4303 and *S. sonnei* 4351.

Annotated genomes were compared with the help of Mauve software (Figure 2) to display genome-wide matches (unidirectional or inverted sections). Mauve software was used to find missing parts and SNPs as well.

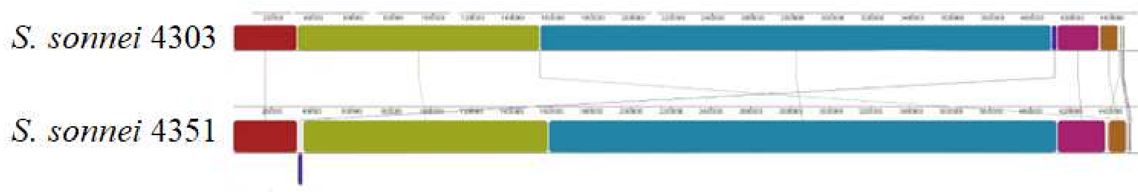

Figure 2. Mauve pairwise comparison of the genomes of *S. sonnei* 4303 and *S. sonnei* 4351.

Plasmid C was found in both *S. sonnei* 4303 and 4351, while plasmid A (virulence plasmid) and B (colicin plasmid) are missing in these strains, supporting the previous data, as *S. sonnei* 4303 was generated from the highly pathogenic phase I *S. sonnei* by plasmid loss during passages.

Genome-wide analysis revealed 256 SNPs, including homopolymer mutations.

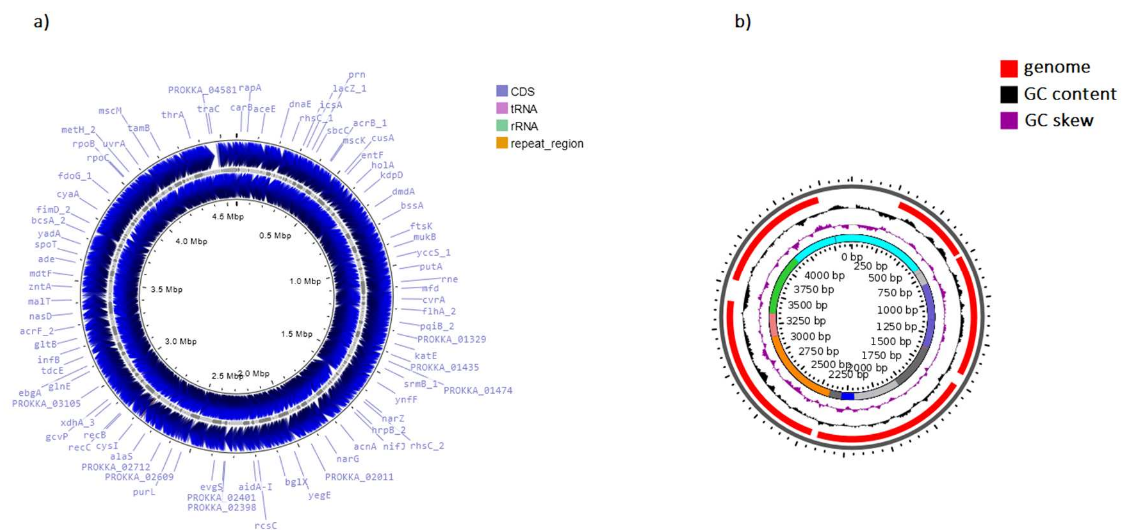

Figure 3. Genomic maps of *S. sonnei* 4351 chromosome according to a) Proksee and b) GView.

Clusters of Orthologous Genes (COG) classification analysis was also performed, that shows the composition of the genome, where each category collects the descendants from the same gene in the ancestral genome.

Clusters of Orthologous Genes	% of genes	% of basepairs
[J] Translation, ribosomal structure and biogenesis	4,51	4,35
[K] Transcription	6,96	6,72
[L] Replication, recombination and repair	4,77	4,60
[D] Cell cycle control, cell division, chromosome partitioning	0,77	0,75
[V] Defense mechanisms	0,95	0,92
[T] Signal transduction mechanisms	4,12	3,98
[M] Cell wall/membrane/envelope biogenesis	5,39	5,20
[N] Cell motility	2,60	2,51

[U] Intracellular trafficking, secretion, and vesicular transport	3,02	2,91
[O] Post-translational modification, protein turnover, and chaperones	3,51	3,38
[C] Energy production and conversion	6,83	6,59
[G] Carbohydrate transport and metabolism	8,66	8,36
[E] Amino acid transport and metabolism	10,03	9,68
[F] Nucleotide transport and metabolism	2,16	2,09
[H] Coenzyme transport and metabolism	3,74	3,61
[I] Lipid transport and metabolism	2,37	2,29
[P] Inorganic ion transport and metabolism	6,83	6,59
[Q] Secondary metabolites biosynthesis, transport, and catabolism	2,19	2,11
[R] General function prediction only	12,16	11,74
[S] Function unknown	8,43	11,62

Table 2. Clusters of Orthologous Genes (COG) classification of annotated genes in *Shigella sonnei* 4351.

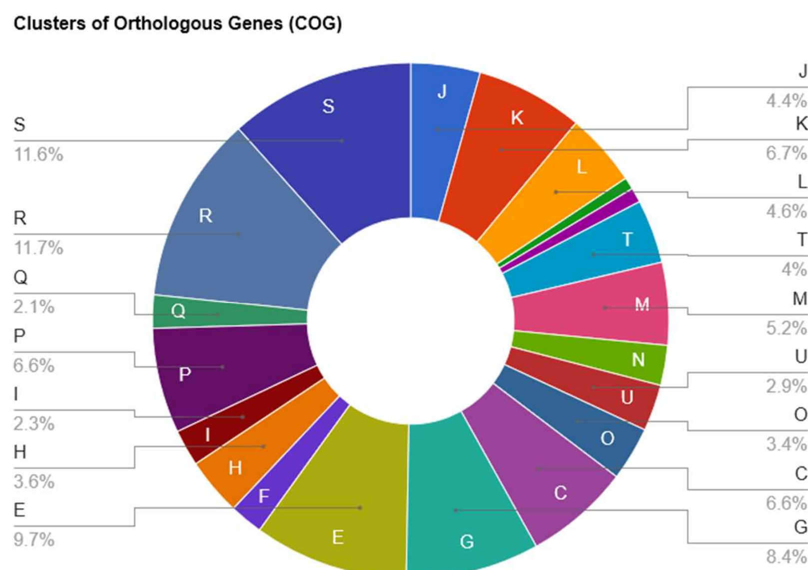

Figure 4. Clusters of Orthologous Genes (COG) classification of annotated genes in Shigella sonnei 4351.

Further comparative genomic analysis of *S. sonnei* 4303, *S. sonnei* 4351 was carried out to analyse lipopolysaccharide biosynthetic genes involving two reference strains, *S. sonnei* Ss046 and *S. sonnei* 53G. This analysis uncovered a mutation in the gene of *rfaD,* also known as *gmhD*. This *rfaD* gene encodes an epimerase enzyme, ADP-L-glycero-D- mannoheptose-6-epimerase required for LPS core biosynthesis.

In a recent publication we propose the connection between the change of LPS structure and the adverse change in the optimal niche. The results show the high significance of the function of RfaD epimerase, and suggest further investigation, as a target, in the fight against Gram-

negative bacterial infections. As part of this quest, the impact of Closantel (a previously described two-component system inhibitor) was tested as an inhibitor of lipopolysaccharide biosynthesis in the *S. sonnei* 4303 strain. By in vitro studies we showed that Closantel has a mild but active inhibitory effect on the expression of the *rfaD* gene, encoding an epimerase enzyme (ADP-L-glycero-D- mannoheptose-6-epimerase) required for LPS core biosynthesis.

The complexes of human serum-, bovine serum- and ovalbumin with bacterial intact lipopolysaccharides or lipooligosaccharides (Escherichia coli, Salmonella enterica, Shigella sonnei, Proteus morganii) were also studied by capillary zone electrophoresis. We investigated the effects of the incubation time, temperature and endotoxin concentrations for the formation of the complex. From the peak position between the protein and the endotoxin-protein complex, we can deduce the chemotypes of endotoxins (lipopolysaccharides or lipooligosaccharides). The amounts of the protein-endotoxin complexes is influenced by the concentration of endotoxins, while the temperature and the incubation time do not have a significant effect on the complex formation.

Knowledge of the type and level of fats in food and clinical matrices is of practical importance, and requires reliable qualitative determination methods. In this regard, GC-MS methods were optimized and validated for the separation of Fatty Acid Methyl Ester (FAME) standard mixtures of 53 components, including compounds with high structural diversity of C4-C24 chain-length. Three capillary columns, about 30 m long, were compared in a novel pairing, two polysiloxane-based commercial columns (the non-polar HP-5MS and the medium/high polarity DB-225MS) and an ionic liquid (IL) - based, extremely polar SLB-IL111 column. Significant differences in the FAME separation properties were observed between the two polysiloxane-based columns and when these columns were compared with the IL-based column. The benefits of utilizing each column were emphasized and by deducting general separation principles, we can predict the separation properties of other isomers. For the separation of cis/trans isomers, the column with the ionic liquid phase proved to be the best; thereby the special interactions with the IL stationary phase were studied. In the light of these findings, it can be said that a sensitive and accurate determination can be performed on the 30 m long ionic liquid-based column, and the results allows the determination of the fatty acid composition of bacterial endotoxins.

Different GC columns were tested and validated for a variety of fatty acyl chains. The extremely polar, ionic liquid-based SLB-IL111 column provided excellent separation properties for the separation of cis and trans fatty acid standards.

The results of IL-based GC column facilitate the precise identification of various types of fatty acids in real samples, e.g., bacteria and endotoxins. For this task, we have optimized the sample preparation of FAME (fatty acid methyl esters), even for hydroxylated fatty acids present in the endotoxins of Gram-negative bacteria. About 50 bacteria cultivated at three different temperatures will be used for FAME analysis.

We have developed an effective method for the separation of basic proteins by capillary zone electrophoresis, using 1,3-dialkylimidazolium-based ionic liquids for dynamic coating and also as background electrolyte components. Excellent separations were obtained for cytochrome c, lysozyme, myoglobin, trypsin, and apo-transferrin, as well as real biological samples (such as proteins from egg whites and human tears), with 100mM ionic-liquid solutions applying 350 V/cm field strength. This method could be used for the interaction studies between different serum proteins and endotoxins.