TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. D OSA - TOM. 927

MEDICA - ODONTOLOGICA

BIOSYNTHESIS OF *YERSINIA ENTEROCOLITICA* SEROTYPE O:3 LIPOPOLYSACCHARIDE OUTER CORE

by

Elise Pinta

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2010

From the Department of Medical Biochemistry and Genetics, University of Turku,
the Infection Biology Research Program, the Department of Bacteriology and
Immunology, Haartman Institute, University of Helsinki
and the Glycoscience Graduate School, Finland

Supervised by

Professor Mikael Skurnik, Ph.D.
Department of Bacteriology and Immunology, the Infection Biology Research
Program, Haartman Institute, University of Helsinki, and
Helsinki University Central Hospital Laboratory Diagnostics, Helsinki, Finland

and

Professor José Antonio Bengoechea, Ph.D.
Program Infection and Immunity, Fundacion Caubet-CIMERA Illes Balears, and
Consejo Superior Investigaciones Científicas (CSIC), Bunyola, Spain

Scientifically Reviewed by

Professor Miguel Valvano, Ph.D.
Department of Microbiology and Immunology, University of Western Ontario, London,
Ontario, Canada

and

Professor Elke Schweda, Ph.D.
Karolinska Institutet, Clinical Research Center, Huddinge Hospital, NOVUM,
Huddinge, Sweden

English language Reviewed by

Professor Joseph S. Lam, Ph.D.
Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario,
Canada

Opponent

Uwe Mamat, Ph.D.
Division of Structural Biochemistry, Research Center, Borstel, Leibniz-Center for
Medicine and Biosciences, Borstel, Germany

Cover picture: Hugo Pinta
ISBN 978-951-29-4429-3 (PRINT)
ISBN 978-951-29-4430-9 (PDF)
ISSN 0355-9483
Painosalama Oy – Turku, Finland 2010

To my family

Elise Pinta

Biosynthesis of *Yersinia enterocolitica* serotype O:3 lipopolysaccharide outer core

Department of Medical Biochemistry and Genetics, University of Turku, Turku, the Infection Biology Research Program, Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki, Helsinki and the Glycoscience Graduate School, Finland
Annales Universitatis Turkuensis
Painosalama, Turku 2010

ABSTRACT

Lipopolysacharide (LPS) present on the outer leaflet of Gram-negative bacteria is important for the adaptation of the bacteria to the environment. Structurally, LPS can be divided into three parts: lipid A, core and O-polysaccharide (OPS). OPS is the outermost and also the most diverse moiety. When OPS is composed of identical sugar residues it is called homopolymeric and when it is composed of repeating units of oligosaccharides it is called heteropolymeric.

Bacteria synthesize LPS at the inner membrane via two separate pathways, Lipid A-core via one and OPS via the other. These are ligated together in the periplasmic space and the completed LPS molecule is translocated to the surface of the bacteria. The genes directing the OPS biosynthesis are often clustered and the clusters directing the biosynthesis of heteropolymeric OPS often contain genes for i) the biosynthesis of required NDP-sugar precursors, ii) glycosyltransferases needed to build up the repeating unit, iii) translocation of the completed O-unit to the periplasmic side of the inner membrane (flippase) and iv) polymerization of the repeating units to complete OPS.

The aim of this thesis was to characterize the biosynthesis of the outer core (OC) of *Yersinia enterocolitica* serotype O:3 (YeO3). *Y. enterocolitica* is a member of the Gram-negative *Yersinia* genus and it causes diarrhea followed sometimes by reactive arthritis. The chemical structure of the OC and the nucleotide sequence of the gene cluster directing its biosynthesis were already known; however, no experimental evidence had been provided for the predicted functions of the gene products. The hypothesis was that the OC biosynthesis would follow the pathway described for heteropolymeric OPS, i.e. a Wzy-dependent pathway.

In this work the biochemical activities of two enzymes involved in the NDP-sugar biosynthesis was established. Gne was determined to be a UDP-*N*-acetylglucosamine-4-epimerase catalyzing the conversion of UDP-GlcNAc to UDP-GalNAc and WbcP was shown to be a UDP-GlcNAc-4,6-dehydratase catalyzing the reaction that converts UDP-GlcNAc to a rare UDP-2-acetamido-2,6-dideoxy-d-*xylo*-hex-4-ulopyranose (UDP-Sug*p*).

In this work, the linkage specificities and the order in which the different glycosyltransferases build up the OC onto the lipid carrier were also investigated. In addition, by using a site-directed mutagenesis approach the catalytically important amino acids of Gne and two of the characterized glycosyltranferases were identified. Also evidence to show the enzymes involved in the ligations of OC and OPS to the lipid A inner core was provided.

The importance of the OC to the physiology of *Y. enterocolitica* O:3 was defined by determining the minimum requirements for the OC to be recognized by a bacteriophage, bacteriocin and monoclonal antibody. The biological importance of the rare keto sugar (Sug*p*) was also shown.

As a conclusion this work provides an extensive overview of the biosynthesis of YeO3 OC as it provides a substantial amount of information of the stepwise and coordinated synthesis of the *Ye* O:3 OC hexasaccharide and detailed information of its properties as a receptor.

Key words: biosynthesis, core oligosaccharide, lipopolysaccharide, *Yersinia*

Elise Pinta

Lipopolysakkaridin ulomman ydinsokerin biosynteesi *Yersinia enterocolitica* (serotyyppi O:3) bakteerissa

Lääketieteellinen biokemia ja genetiikka, Turun yliopisto, Infektiobiologian tutkimusohjelma, Bakteriologian ja immunologian osasto, Haartman instituutti, Helsingin yliopisto, sekä valtakunnallinen Glykotieteiden tutkijakoulu
Annales Universitatis Turkuensis
Painosalama, Turku 2010

TIIVISTELMÄ

Lipopolysakkaridi (LPS) on osa bakteerin soluseinää. Bakteerien pintarakenteet, LPS mukaan lukien, ovat hyvin monimuotoisia ja tärkeitä bakteerien sopeutumiselle.

LPS makromolekyylin hydrofobinen pää ujuttautuu ulompaan ulkokalvon lipidikerrokseen ja hydrofiilinen pää työntyy bakteerin ympäristöön. Rakenteellisesti LPS jaetaan kolmeen osaan: Lipidi A, ydinsokeri ja oligopolysakkaridi (OPS), joista OPS uloimpana rakenteena on myös monimuotoisin. OPS voi koostua joko yksittäisistä sokeritähteistä (nk. homopolymeerinen), tai useamman sokerin muodostamista toistuvista yksiköistä (nk. heteropolymeerinen).

Bakteeri muodostaa LPS:n sisäkalvollaan kahdessa osassa: lipidi A ydinsokerin yhtenä ja OPS:n toisena yksikkönä. Vasta valmis LPS molekyyli siirretään bakteerin ulkopinnalle. Geenit, jotka ohjaavat OPS:n biosynteesiä, ovat usein ryhmittyneet yhteen geeniryhmiksi. Heteropolymeerisen OPS:n syntyä ohjaavasta geeniryhmästä löytyvät yleensä geenit i) tarvittavien sokeritähteiden synteesiin (nk. nukleotidisokereita syntetisoivat entsyymit) ii) sokeritähteiden toisiinsa liittämiseen (glykosyylitransferaasit), iii) valmiin O-yksikön siirtämiseen sisäkalvon periplasmiselle puolelle (flippaasi) ja iv) yksittäisten O-yksiköiden liittämiseen valmiiksi OPS:ksi.

Tämän työn tarkoituksena oli selvittää miten LPS:n ulompi ydinsokeri syntyy *Yersinia enterocolitica* serotyyppi O:3 -bakteerissa. *Y. enterocolitica* on gram-negatiivinen *Yersinia* suvun bakteeri, joka aiheuttaa vatsatautia ja jälkitautina toisinaan nivelkipuja. Entuudestaan serotyyppi O:3:n ulomman ydinsokerin tiedettiin olevan tärkeä bakteerin taudinaiheuttamiskyvylle. Ulomman ydinsokerin kemiallinen rakenne oli selvitetty ja sen biosynteesiä ohjaavan geeniryhmän emäsjärjestys tunnettiin. Kullekin geenille oli emäsjärjestyksen perusteella ennustettu tehtävä mutta yhdenkään geenituotteen toimintaa ei oltu kokeellisesti todistettu. Ulomman ydinsokerin biosynteesin oli ennustettu seuraavan heteropolymeerisille OPS:lle tyypillistä tapaa.

Tässä työssä osoitettiin kahden nukleotidisokerien biosynteesiin liittyvän entsyymin aktiivisuus: Gne katalysoi UDP-*N*-asetylglukosamiinin (UDP-GlcNAc) nelos hiilen (C4) hydroksyyliryhmän epimerisaatiota tuottaen UDP-*N*-asetyyliglalaktosamiinia ja WbcP on 4,6-dehydrataasi katalysoiden reaktiota, jossa UDP-GlcNAc:ista syntyy harvinaista UDP-2-acetamido-2,6-dideoxy-d-*xylo*-hex-4-ulopyranoosia (UDP-Sug*p*).

Tässä työssä selvitettiin myös keskinäinen järjestys ja tarkat sidokset, joilla kuusi eri glykosyylitransferaasia yksitellen lisäävät oikeanlaiset sokeritähteet lipidikantajan päälle valmiiksi ulommaksi ydinsokeriksi.

Muuttamalla kohdistetusti Gne entsyymin ja kahden glykosyylitransferaasin aminohappoja selvitetiin molekyylitasolla kullekin entsyymiaktiivisuudelle tärkeitä rakenteita. Tässä työssä löydettiin myös entsyymit, jotka liittävät ulomman ydinsokerin ja OPS:n sisempään ydinsokeriin.

Lisäksi ulomman ydinsokerin biologista merkitystä tutkittiin selvittämällä sokerirakenteet joihin ydinsokerintunnistavat bakteriofagi, bakteriosiini ja monoklonaalinen vasta-aine sitoutuvat. Myös ulommasta ydinsokerista löytyneen harvinaisen ketosokerin osoitettiin olevan biologisesti tärkeä.

Avainsanat: biosynteesi, lipopolysakkaridi, ydinsokeri, *Yersinia*

Contents

ABBREVIATIONS

ABC-2	ATP-binding cassette-2
Ac	acetyl
ACP	acyl carrier protein
ADP	adenine diphosphate
Alt	altrose
AU	activity units
CDP	cytosine diphosphate
CE	capillary elctrophoresis
CFU	colony forming units
CIN	cefsulodin-irgasan-novobiocin
DIG	digoxigenin
DOC	deoxycholate
E. coli	*Escherichia coli*
ECA	enterobacterial common antigen
f	furanose sugar configuration
Fo	Formyl
Fuc	fucose
FucNAc	*N*-acetylfucosamine
Gal	galactose
GalNAc	*N*-acetylgalactosamine
G+C	Guanine + Cytosine
GDP	guanine diphosphate
Glc	glucose
GlcN	glucosamine
GlcNAc	*N*-acetylglucosamine
GTase	glycosyltransferase
Gul	gulose
Hep	L-*glycero*-D-*manno*-heptose
IC	inner core
IM	inner membrane
JUMPStart	Just Upstream of many Polysaccharide associated gene starts
kb	kilo-base pair(s)
Kdo	3-deoxy-D-*manno*-octulosonic acid
LA	Luria Bertani agar
L-Ara4N	4-amino-4-deoxy-L-arabinose
LB	Luria Bertani medium
LPS	lipopolysaccharide
mAb	monoclonal antibody
Man	mannose
me	methyl

$NAD(P)^+$	nicotinamide adenine dinucleotide phosphate
NC-IUBMB	Nomenclature Committee of the International Union of Biochemistry and Molecular Biology
NDP	nucleoside diphospate
NMP	nucleoside monophosphate
O-ag	O-antigen, O chain
OC	outer core
OM	outer membrane
ops	operon polarity suppressor
OPS	O- polysaccharide (O-ag)
P.	*Pseudomonas*
p	pyranose sugar configuration
PAGE	polyacrylamide gel electrophoresis
*P*Cho	phosphorylcholine
PCP	polysaccharide co-polymerase
PCR	polymerase chain reaction
pfu	plaque forming units
PHPT	polyprenyl-phosphate hexose-1-phosphate transferase
PmB	polymyxin B
PNPT	polyprenyl-phosphate *N*-acetylhexosamine-1-phosphate transferase
PPEtN	pyrophosporylethanolamine
PST	polysaccharide-specific transport
QuiNAc	*N*-acetylquinovosamine, 2-acetamido-2,6-dideoxy-d-glucose
R	rough
RT	room temperature (20 ºC to 25 ºC)
S	smooth
SDR	short-chain dehydrogenase/reductase
SDS	sodium dodecyl sulphate
SR	semi-rough
Sug*p*	2-acetamido-2,6-dideoxy-D-*xylo*-hex-4-ulopyranose
TLR4/MD-2	Toll-like receptor 4 complexed with MD-2 (MD-2 is a name for gene / protein, not an abbreviation) (Burk & Klempnauer, 1991, Kim et al, 2007)
TSB	tryptic soy broth
UDP	uridine diphosphate
UDP-GalNAcA	UDP-*N*-acetyl-D-galactosaminuronic acid
Und-P	undecaprenyl phosphate
Y.	*Yersinia*
Y. e.	*Yersinia enterocolitica*
Y. p.	*Yersinia pestis*
Y. pstb.	*Yersinia pseudotuberculosis*
YeO3	*Yersinia enterocolitica* serotype O:3
YeO8	*Yersinia enterocolitica* serotype O:8

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original publications, which are referred to in the text by the Roman numerals (I-IV):

I. Bengoechea, J. A., **Pinta, E**., Salminen, T., Oertelt, C., Holst, O., Radziejewska-Lebrecht, J., Piotrowska-Seget, Z., Venho, R., and Skurnik, M. (2002) Functional characterization of Gne (UDP-N-acetylglucosamine-4- epimerase), Wzz (chain length determinant), and Wzy (O-antigen polymerase) of *Yersinia enterocolitica* serotype O:8. *J Bacteriol* **184**: 4277-4287.

II. **Pinta, E.**[1], Duda, K.A.[1], Hanuszkiewicz, A.[1], Kaczyński, Z., Lindner, B., Miller, W., Hyytiäinen, H., Vogel, C., Borowski, S., Kasperkiewicz K., Lam, J.S., Radziejewska-Lebrecht, J., Skurnik, M., and Holst, O., (2009) Identification and role of a 6-deoxy-4-keto-hexosamine in the lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3. *Chemistry*. 2009 Sep 28;15(38):9747-54.
[1] These authors contributed equally to this work

III. **Pinta, E.**, Duda, K.A., Hanuszkiewicz, A., Salminen, T.A., Bengoechea, J.A., Hyytiäinen, H., Lindner, B., Radziejewska-Lebrecht, J., Holst, O., and Skurnik, M., Characterization of the six glycosyltransferases involved in the biosynthesis of *Yersinia enterocolitica* serotype O:3 lipopolysaccharide outer core. *J. Biol. Chem*. 2010 Sep 3;285(36):28333-42

IV. **Pinta, E.,** Li, Z., Batzilla, J., Pajunen, M., Rakin, A., and Skurnik, M. Identification of *Yersinia enterocolitica* serotype O:3 lipopolysaccharide ligases. Manuscript in preparation.

1. INTRODUCTION

Genus *Yersinia* consists of 14 species among which *Y. pestis, Y. pseudotuberculosis* and *Y. enterocolitica* are human pathogens. *Y. pestis* causes bubonic plague while *Y. pseudotuberculosis* and *Y. enterocolitica* cause yersiniosis, a diarrheal disease which is sometimes followed by post-infectious reactive arthritis.

Yersinia lipopolysaccharide (LPS) is a surface structure encountering the surrounding environment of the bacteria and therefore it is not surprising that LPS structures are critical for virulence. LPS has three main structural moieties: lipid A (two glucosamine residues substituted with fatty acid chains and phosphate groups); core (an oligosaccharide containing about ten to fifteen sugar residues) and O side chain (O-antigen, OPS, a polysaccharide chain composed of repeat units of one to eight sugar residues).

LPS biosynthesis proceeds as two separate processes; lipid A-core as one and the OPS (synthesized on a carrier lipid) as another. The OPS is ligated to lipid-A core at the periplasmic side of the inner membrane and the LPS in then translocated to the outer membrane.

In this thesis research, my initial aim was to attain a comprehensive understanding of the genetics, biochemistry and biological role of LPS in *Yersinia*. More specifically, this project focused mainly on the biosynthesis of *Y. enterocolitica* serotype O:3 (YeO3) outer core (OC) and, where similar, on the biosynthesis *Y. enterocolitica* serotype O:8 (YeO8) OPS. Both YeO3 OC and YeO8 OPS have been shown to be virulence factors. They are required during the early phases of infection and OC mutant bacteria are sensitive to cationic antimicrobial peptides. My hypothesis was that the YeO3 OC biosynthesis would follow the pathway described for heteropolymeric OPS, i.e. a Wzy-dependent pathway.

In this work, I determined the functions of nine genes needed for the biosynthesis of the outer core oligosaccharide of the *Y. enterocolitica* serotype O:3 LPS of which some are shared with *Y. enterocolitica* O:8. This included the characterization of three different types of enzymes: biosynthetic enzymes needed for the synthesis of NDP-sugars, glycosyltransferases and ligases.

2. REVIEW OF LITERATURE

2.1. Genus Yersinia

The *Yersiniae* are Gram-negative, small, rod-shaped bacteria. The genus *Yersinia* consists of 14 species among which *Yersinia pestis (Y. p.), Yersinia pseudotuberculosis (Y. pstb.)* and *Yersinia enterocolitica* (*Y. e.)* are human pathogens (Butler, 1983, Merhej et al, 2008, Sprague & Neubauer, 2005, Sprague et al, 2008). *Y. p.* causes bubonic plague while *Y. pstb.* and *Y. e.* cause mostly food-borne yersiniosis, usually a diarrheal disease sometimes followed by post-infectious reactive arthritis. (Carniel et al, 2006)

All pathogenic *Yersinia* species have a 70-kb virulence plasmid that carries the genes coding for many essential virulence factors. However, many virulence genes are also located in the chromosome. Only the genomic sequences of *Y. p., Y. pstb.* serotypes O:1 and O:3, and *Y. e.* serotype O:8 were available until recently when the genomic sequences of eight other *Yersinia* species were published (Chen et al, 2010). By combining the sequence information from the new databases Chen *et al.* were able to determine the core genome encoding about 2,500 proteins that define the *Yersinia* genus. Sequencing of the genome of the three remaining *Yersinia* species is under way.

2.1.1. Y. enterocolitica

Y. enterocolitica is widely distributed in nature, in aquatic and animal reservoirs. Swine is apparently a major reservoir for human pathogenic strains (Bottone, 1997).

Since *Y. e.* is heterogenous species, originally, for the need of appreciation of the pathogenic potential of a given isolate, *Y. e.* has been separated into six biotypes. They are based predominantly on biochemical reactions. The biogrouping correlates also with serogroup designation, human pathogenicity and ecologic and geographic distribution (Bottone, 1999). Isolates belonging to biogroup 1A are salicin-, esculin- and pyrazinamidase-positive and regarded as avirulent; while isolates from biogroups 1B, 2, 3, 4 and 5 are associated with human infections.

Serologically, *Y. e.* and *Y. e.* like bacteria can be separated into 60-70 serogroups of which 11 have been most frequently associated with human infections (Bottone, 1999, Carniel et al, 2006). Serogroups O:3, O:5,27, O:8 and O:9 belonging to biogroups 4, 2, 1B and 2, respectively, are the most clinically significant (Carniel et al, 2006).

In addition to the above mentioned traditional typing methods that rely on different phenotypes, bacterial strains are also grouped using genetic methods. Several different methods have been developed and the emerging nucleotide sequence data helps to develop the grouping further. The molecular-based typing methods fall into three general categories: (i) those based on restriction analysis of the bacterial DNA, (ii) those

based on PCR amplification of particular genetic targets; and (iii) those based on the identification of DNA sequence polymorphism (Foley et al, 2009). In the case of *Y. e.*, the grouping of strains based on genotyping was in good agreement with grouping using the previous typing systems. Six biogroups were established in Y. e. (Boghenbor et al, 2006, Thoerner et al, 2003). In addition to biotyping, genotyping allows the comparison between bacterial isolates in more details. For instance, the use of genotyping has shown that biotypes 2, 3, and 4 are more closely related to each other than to 1A and 1B; and biotype 1A exhibits the greatest genetic heterogeneity. (Boghenbor et al, 2006)

2.1.2. Y. pestis

Y. pestis is an obligate parasite and the etiologic agent of plague. Although bubonic and pneumonic plagues are serious human diseases, the human host plays no role in the long term survival of *Y. p.* because plague is a zoonotic disease affecting rodents as its primary host (Perry & Fetherston, 1997). Transmission between rodents is generally accomplished by fleas carrying these organisms.

Traditionally *Y. p.* are separated into three biotypes (also called biovars) according to their abilities to convert nitrate to nitrite and ferment glycerol (Devignat, 1951). More recently, a fourth biovar has been proposed based on glycerol and nitrate reactions, as well as the inability to ferment arabinose (Zhou et al, 2004). Strains of the different classical biotypes (I-III) are known to be virulent for animals or humans (Perry & Fetherston, 1997), however, the representative strain of the fourth biotype (91001) was found to be avirulent (Zhou et al, 2004).

Y. p. has typical cell envelope and expresses enterobacterial common antigen (ECA) as other enteric bacteria but due to the lack of O-antigen *Y. p.* strains are serologically very homogenous (Brubaker, 1991). Genotypically, *Y. p.* has been divided into 8 major genetic groups. The genetic grouping is only partially compatible with the classical typing to biovars. (Achtman et al, 2004)

2.1.3. Y. pseudotuberculosis

Y. pseudotuberculosis is widely spread in the environment (soil and water) and it is a common inhabitant of the intestine in a wide variety of domestic and wild animals (Fredriksson-Ahomaa, 2007).

Y. pstb. isolates have been classified into four biotypes (1, 2, 3 and 4) according to their biochemical properties in the usage of citrate, melibiose and raffinose. There is no clear correlation of pathogenicity with the biotype, but melibiose-positive strains (biotypes 1 and 4) have been shown to be more pathogenic than melibiose-negative strains (biotypes 2 and 3 respectively). (Fredriksson-Ahomaa, 2007)

There are 21 different serotypes (O:1a, 1b, 1c, 2a, 2b, 2c, 3, 4a, 4b, 5a, 5b, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15) and genetically the *Y. pstb.* species seem to be a more homogenous

group than *Y. e.* (Carniel et al, 2006). The serotyping scheme of *Y. pstb.* is based on serological tests whereby antisera recognize the differences in the O antigens among the strains in this species. More recently an O-genotyping method has been developed (Bogdanovich et al, 2003). The genotyping method was capable of identifying a strain belonging to a certain genotype, even when it has lost the ability to express O-antigen and thus could not be serotyped.

It is also worth mentioning that *Y. p.* and *Y. pstb.* are more closely related to each other than to *Y. e.* or other *Yersinia* species. In fact there were evidence that *Y. p.* had quite recently evolved from *Y. pstb.* serotype O:1b (Achtman et al, 1999, Skurnik et al, 2000). And by O-genotyping it is also possible to distinguish *Y. p.* from *Y. pstb.* (Bogdanovich et al, 2003).

2.2. Lipopolysaccharide

Lipopolysaccaharide (LPS) is found on the outer leaflet of Gram-negative bacteria. They are anchored by their lipid A region to the outer lipid layer of the outer membrane (OM) with the proximal hydrophilic sugar of the O-Ag facing the surrounding environment. LPS is a major component of the OM and it is crucial for bacterial survival (Gronow et al, 2010, Raetz & Whitfield, 2002, Raetz et al, 2007).

LPS can be divided into two structurally different moieties. One is the lipid containing moiety that is called lipid A and the other is the carbohydrate constituent. The latter can be sub-divided further, and the distal part is called O-antigen or O-polysaccharide (O-Ag, O-chain, OPS) and the middle region of the LPS that interconnects lipid A to OPS is called core oligosaccharide (Raetz & Whitfield, 2002, Raetz et al, 2007). Lipid A is generally essential for the biosynthesis of LPS and the core oligosaccharide is invariably also important, but the OPS on the other hand can be missing. (Gronow et al, 2010)

Lipid A is composed of two glucosamine (GlcN) residues substituted with fatty acid chains. This is the most conserved region of the LPS. Lipid A is responsible for the endotoxic effect and it is therefore called an endotoxin (Raetz et al, 2007).

The core is an oligosaccharide containing about ten to fifteen sugar residues. It is moderately diverse and sometimes it can be further divided to inner and outer core (IC and OC, respectively). The main elements of the IC are highly conserved whereas the OC is rather variable (Gronow et al, 2010). Together with lipid A, the core affects OM stability; mutations affecting lipid A structure are usually lethal to the organism while core mutants increase the permeability of OM to hydrophobic compounds.

The outermost and highly variable region of LPS is the OPS. OPS is a carbohydrate chain consisting of repeating units of one to eight monosaccharide units (Knirel & Kochetkov, 1994, Knirel, 2010). Constituents of OPS range from sugars that are widespread in nature to some that are rare. Some of the sugars are highly uncommon and have not been found in other natural carbohydrates (Knirel & Kochetkov, 1994, Knirel, 2010). The

OPS where the repeating unit is a monosaccharide are called homopolymeric and where the repeating unit is an oligosaccharide of different sugars are called heteropolymeric, respectively. Homo- and heteropolymeric OPS can be easily distinguished from one another by polyacrylamide gel electrophoresis (PAGE, Figure 1), where high molecular weight polymers of LPS with heteropolymeric OPS gives a ladder like appearance and the one with homopolymeric OPS looks more like a smear due to the fact that the PAGE can not resolve it to separate bands (Figure 1). Between bacterial strains the OPS varies considerably and is often the basis of the serotyping of Gram-negative species (Knirel & Kochetkov, 1994, Knirel, 2010).

The strains producing OPS are called smooth (S) as a consequence of their smooth colony morphology. The strains that lack OPS have rough-looking colony morphology and are thus called rough (R). Strains in which their LPS has only one O-unit are called semi-rough (SR).

In most bacteria the genes directing the biosynthesis of core and OPS form distinct clusters that map to different parts of the bacterial chromosome. In contrast, the genes encoding lipid A biosynthesis are scattered around the genome (Cunneen et al, 2009).

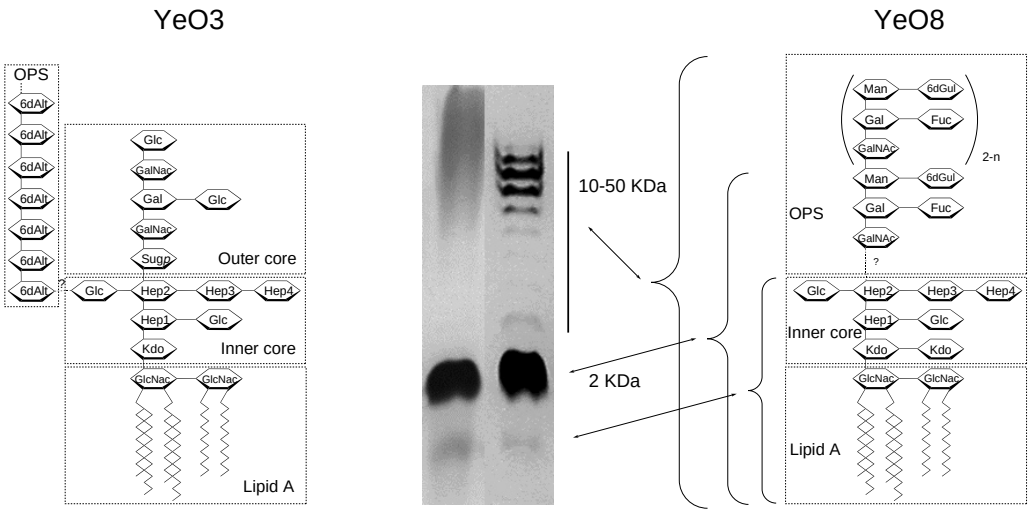

Figure 1. General assembly of LPS with homopolymeric or heteropolymeric OPS. As examples LPS of *Y. e.* serotypes O:3 (YeO3) and O:8 (YeO8) analyzed on silver-stained DOC-PAGE.

2.2.1. Saccharide structures and nomenclature

The enormous diversity provided by different saccharide units becomes obvious when discussing LPS. To help the reader to follow the structures, a brief overview of the nomenclature is provided here. As the systematic names according to organic chemistry nomenclature are often long and impractical in the case of naming mono- and especially polysaccharide structures like core and O-chain, some of the different monosaccharide

isomers and disaccharides have been named with trivial names. For example general formula $C_6H_{12}O_6$ has 4 asymmetric carbons (carbons 2, 3, 4, and 5) and can exist in any one of the 16 possible isomeric forms. The different forms (easily demonstrated by Fisher projections arising from the different orientations of hydroxyl groups around carbons 2, 3, and 4) have been named talose, galactose, idose, gulose, mannose, glucose, altrose and allose and as all of them can be either L or D forms it leads to above-mentioned 16 isomeric forms. A monosaccharide is a member of the D-series if the hydroxyl group on the chiral carbon farthest from carbon 1 (in our example carbon 5) is also on the right in the Fisher projection. Unless specified otherwise the D configuration (which predominates in nature) is implied in these abbreviated names. In addition for simple monosaccharides three-letter abbreviations are frequently used (e.g. Gal for galactose and Glc for glucose). In solution, sugar molecules exist as cyclic hemiacetals and hemicetals, and for chemical stability reasons, five- and six-membered rings are the most common. The cyclic form generates a new asymmetric carbon called anomeric carbon and if its hydroxyl group is directed to the same side as the hydroxyl group at the highest numbered asymmetric carbon, the sugar is termed α (in opposite situation, β). A five-membered hemiacetal ring is called a furanose, while a six-membered hemiacetal ring is called a pyranose. The ring sizes may be included with the abbreviated name of the monosaccharide as *f* or *p*, for example Gal*f* or Glc*p*.

One sugar is ubiquitously present in every LPS structure investigated so far, i.e. 3-deoxy-*α*-D-*manno*-oct-2-ulosonic acid (Kdo), and a second highly conserved sugar is L-*glycero*-D-*manno*-heptose (Hep) (Gronow et al, 2010). Both of these are found as parts of the IC structures but when taken into account the more diverse parts of LPS (OS and OPS) all together more than 60 different monosaccharide units have been identified (Knirel & Kochetkov, 1994). To clarify the classification the aminosugars (like 4-amino-4-deoxy-L-arabinose (L-Ara4N) and monosaccharides having carboxyl groups (like hexuronic acid D-GlcA and D-GalA) as well as monosaccharides having both amino and carboxyl group (like 2-amino-2-deoxy-hexuronic acid D-ManNA) are counted here as different monosaccharides. To these monosaccharides non carbohydrate groups can be added as substituents (like acetyl group in 2-acetylamino-2-deoxy-D-glucose (*N*-acetyl-D-glucosamine, GlcNAc)) and they are discussed later. (Knirel & Kochetkov, 1994, Knirel, 2010)

2.2.2. Fatty acid structures and nomenclature

In addition to the different monosaccharides LPS structures vary also in their fatty acid composition of the lipid A. As a characteristic feature, lipid A contains saturated medium- to long-chain (10-22 carbon atoms) fatty acids in ester and amide linkages. These fatty acids can be nonhydroxylated or when hydroxylated, the 3-hydroxylated fatty acids preponderate. When fatty acids are linked directly to hydroxyl or amino groups of the lipid A backbone the acyl groups are termed *primary* acyl groups. These include 3-hydroxy and 3-keto fatty acids. When the fatty acids are linked to the 3-hydroxyl groups of the ester or amide-linked 3-hydroxy fatty acids they are termed *secondary*

acyl residues (Zahringer et al, 1994). The most common fatty acids have trivial names in addition to the systematic name. In Table 1 are listed some trivial and systematic names (1A) as well as some examples of acyl and 3-hydroxyacyl residues of lipid A backbone of various bacteria (1B).

Table 1. A) trivial and systematic names of some fatty acids generally found in LPS structures. B) examples of acyl and 3-hydroxyacyl residues of lipid A backbone of various bacteria (Oertelt et al, 2001, Zahringer et al, 1994).

A

Number of carbons / double bonds	Trivial name	Systematic name	Formula
12 / 0	Lauric acid	n-dodecanoic acid	$CH_3(CH_2)_{10}COOH$
14 / 0	Myristic acid	n-tetradecoic acid	$CH_3(CH_2)_{12}COOH$
16 / 0	Palmitic acid	n-hexadecoic acid	$CH_3(CH_2)_{14}COOH$
18 / 0	Stearic acid	n-oktadecoic acid	$CH_3(CH_2)_{16}COOH$

B

Bacteria	GlcN(II) at position		GlcN(I) at position	
	3´(ester)	2´amide	3 (ester)	2 (amide)
E. coli	14:0[3-*O*(14:0)]	14:0[3-*O*(12:0)]	14:0(3-OH)	14:0(3-OH)
S. typhimurium	14:0[3-*O*(14:0)]	14:0[3-*O*(12:0)]	14:0(3-OH)	14:0(3-OH)
H. influenzae	14:0[3-*O*(14:0)]	14:0[3-*O*(14:0)]	14:0(3-OH)	14:0(3-OH)
P. aeruginosa	10:0(3-OH)	12:0[3-*O*(12:0)] 12:0[3-*O*[12:0(2-OH)]]	10:0(3-OH)	12:0[3-*O*(12:0)] 12:0[3-*O*[12:0(2-OH)]]
Y. e. O:8		14:0[3-*O*(C14:0)]	14:0(3-OH)	14:0(3-OH)

2.2.3. Lipopolysaccharide structures

Lipopolysaccharides demonstrate extreme variability. Gram-negative bacterial species have unique LPS structures and even a single bacterial species can contain various LPS structures as bacteria from different serotypes often have different O-ag structures. For instance, 17 different OPS structures of *Y. pstb.* have been reported (Knirel & Kochetkov, 1994, Knirel, 2010). Until now, thousands of LPS structures from hundreds of strains have been characterized, at least partially, to reveal chemical details and the enormous structural variations have become evident. Updated information of the published structures is available at Bacterial Carbohydrate Structure DataBase (http://www.glyco.ac.ru/bcsdb3/).

All the LPS molecules on the surface of a single bacterium are not uniform, they vary by their length. So a single bacterium carries a population of LPS: all LPS molecules have lipid A and most also intact core, but some do and some do not have the O chain (Figure 1). Also the O-chain length usually varies and the chain length has modal distribution, typical for the analyzed strain.

The expressed LPS on the surface of a single strain can also vary according to different factors. This way the bacteria adapts to the needs of different environments. All the three structural moieties of LPS have been demonstrated to have modifications (Raetz & Whitfield, 2002). Some of the variations have been reported to occur in many bacteria (mainly studied in *Escherichia coli, Salmonella* and *Pseudomonas)* but others seem to be more species specific. As an example, conditions like low Mg^{2+} activate modification of lipid A with palmitate (Guo et al, 1998, Raetz & Whitfield, 2002) (Table 1A). Also mildly acidic conditions affect the lipid-A composition via attachment of phosphoetanolamine and 4-amino-4-deoxy-L-arabinose (L-Ara4N). These modifications have been reported to be beneficial for the survival of the bacteria e.g. the attachment of Ara4N leads to polymyxin B (PmB) resistance in number of bacteria. (Gronow et al, 2010, Nikaido, 2003).

Also, the growth temperature affect the LPS expression; for example, the OPS expression of *Y. e.* serotypes O:3 (YeO3) and O:8 (YeO8) (discussed later), and lipid A composition of *E. coli.* The latter has been noticed to change when the bacteria were grown under cold shock (12 ºC) probably to adjust its outer-membrane fluidity (Carty et al, 1999, Raetz & Whitfield, 2002).

It is worth mentioning that the different modifications are not always stoichiometric, but the LPS extracted from a single strain have heterogeneity in the degree of modification. This is partly because of the above-mentioned diversity of the LPS structural; thus, structural analyses face some challenges in that sometimes only the structure of the predominant carbohydrate backbone is known (e.g. to the IC) and this base structure can be decorated with nonstoichiometric additions of other sugars and additional groups e.g. phosphates. (Raetz & Whitfield, 2002)(see also below).

2.2.3.1. Different groups present in LPS

The diversity of LPS arises owing to several reasons. A wide variety of different monosaccharide units as well as the configurations and positions (different carbons) of the linkages connecting them provide the basis for the diversity. In addition, OPS can be linear or branched and the number of fatty acid chains and their length varies in lipid A (Raetz & Whitfield, 2002). Furthermore, a number of functional groups can be present in LPS, *i.e.,* in the OPS, core and lipid-A the constituent monosaccharides can have functional groups attached to them. For example, *N*-linked acetyl groups in sugars like GlcNAc and 2-acetylamino-2-deoxy-D-galactose (*N*-acetyl-D-galactosamine, GalNAc), but also noncarbohydrate substituents may be attached to sugars by several different linkages. Ether linkage is found in partially-methylated monosaccharides, ester linkage in *O*-acetylated and phosphorylated monosaccharides, and acetal linkage in acetals of pyruvic acid (Knirel & Kochetkov, 1994). Altogether, 30 different non carbohydrate components have thus far been recognized (Raetz & Whitfield, 2002). Examples of the different linkages and components as well as the microorganisms where they are found are shown in Table 2. For more examples see Table 2 in (Knirel & Kochetkov, 1994, Knirel, 2009).

Table 2. Examples of noncarbohydrate components of O-antigens (modified from (Knirel & Kochetkov, 1994)).

Linkage type	Component	Abbreviation	Microorganism
N-Acyl	Formyl	Fo	*E. coli*
(at amino group			*Y. e.*
of amino sugar)	Acetyl	Ac	w
	N-Acetyl-L-seryl	Ac-L-Ser	*E. coli*
O-Alkyl	Methyl	*O-me*	*P. maltophila*
			Campylobacter jejuni
O-Acyl	Acetyl	*O*-Ac	w
Amido	Amino		*P. aeruginosa*
(at carboxyl	L-Seryl	L-Ser	*Proteus mirabilis*
group of uronic acid)			
Phosphodiester	Glycerol	Gro	*E. coli*
	D-Ribitol	D-Rib-ol	*Proteus mirabilis*

w., widespread in O-antigens

The Table 2 above focuses on noncarborydrate components of O-antigens but as mentioned earlier, noncarbohydrate components are also found in deeper parts of LPS; *e.g.* phosphate groups are found both decorating the IC and as part of the lipid A. In addition, pyrophosporylethanolamine (PPEtN) and phosphorylcholine (*P*Cho) can decorate the base structure of IC by nonstoichiometric (Raetz & Whitfield, 2002).

2.2.4. Lipid A core linked polymers

Generally, a given isolate produces one O antigen but some bacteria contain multiple lipid A-core linked polymers, and thus it complicates both the structural analysis and O-antigen terminology. *E. coli* provides good example since its lipid A-core can act as an anchor for OPS, for ECA polymer and for some capsular K (Raetz & Whitfield, 2002, Whitfield & Roberts, 1999). *P. aeruginosa* produces two lipid A-core-linked polymers: B-band LPS corresponding to serotype specific O antigen and A-band LPS analogous to lipid-A linked ECA. (Raetz & Whitfield, 2002, Rocchetta et al, 1999)

2.2.5. The LPS diversity and its importance

The degree of diversity of LPS grows from its anchoring region of the molecule. The structural varieties of lipid A and core are associated with the OM stability and endotoxic effect. OPS is thought to act as a shield and it is required during the early phases of infection.

Phosphate groups are important since the phosphorylation of the IC is thought to be critical for the stability of the OM and the two phosphate groups together with the acyloxyacyl groups of *E. coli* lipid A are needed to trigger full TLR4/MD2 (Toll-like reseptor 4/MD2) activation in human cells (Raetz et al, 2007, Rietschel et al, 1994).

The high diversity of the OPS has been associated with the pressure that bacteria are subjected to during evolution to gain protective mechanisms against the action of bacteriophages specific to surface structures of the cell and against the immune systems of higher animals. OPS is an important component in resistance to serum mediated killing, phagosytosis and killing by cationic peptides (Nesper et al, 2001, Whitfield et al, 1997). In different bacteria, the degree of structural differentiation varies significantly. In some species, all representative strains that have been examined produce the same or structurally similar OPS. In other species the OPS contains characteristic components and has characteristic structural types but there are also species where it is difficult to reveal any common features. On the other hand, among different microorganisms, some of the members that are taxonomically remote from each other could produce similar or even identical polysaccharides and it has been concluded that the O-antigen structure cannot be considered as chemotaxonomic criterion (Knirel & Kochetkov, 1994, Knirel, 2009).

2.2.6. Origin of the diversity

The polymorphism of OPS in bacteria has been for a long time proposed to be maintained by some form of balanced selection (Reeves, 1993). Is seems that the environment and the host immune system together have exerted intense selection pressure on the microbial organisms such that they evolve. Therefore producing different OPS forms within a species becomes a necessity for Gram-negative species to exist in nature (D Souza et al, 2005). Horizontal gene transfer is commonly reported for bacterial genes and the reason behind it probably relates to the very nature of the single-celled organism as the newly transferred gene will be of immediate benefit to the recipient and its progeny (Lan & Reeves, 2000). The horizontal gene transfer has been proposed to be at least partially behind the O-ag polymorphism and indeed the variation in mol% G+C (Guanine + Cytosine) indicates that the OPS gene clusters were assembled from several sources (Reeves, 1993).

The possible events behind the current OPS profiles have been studied by using several different computer-based (*in silico*) sequence-analysis algorithms. Identification of different characteristics of the sequences like palindromic units and bacterial interspersed mosaic elements as well as the differences in mol% G+C of gene blocks and variations in base 3 of codons are used to hypothesize the historical events (Stevenson et al, 2000). For instance, identification of symmetrical putative recombination sites have provided evidence for the predictions of the source of putative 6d-L-Alt*f* (altrose) biosynthesis genes (*ddhB-altA-altB)* in *Y. pstb.* O:11 (Cunneen et al, 2009).

The OPS gene clusters appear to have a mosaic structure, genes involved in biosynthesis of a certain nucleoside diphosphate (NDP)-sugar and the corresponding glycosyltransferase (GTase) are often grouped together and thus more easily transferred as one unit from strain to strain.

In some bacterial species, certain OPS genes are not chromosomal but localized in plasmids (e.g. *wbbP* in *Shigella dysenteriae)* and there are studies proposing that the origin of the OPS genes found in plasmids would be in chromosomes of other strains. The above mentioned *wbbP* is thought to be transferred from another species in the human gastrointestinal tract (Feng et al, 2007). New sub-serotypes have also been proposed to be arisen from recombination of OPS clusters. As a case in point, the generation of *E. coli* O9a the original cluster of *E. coli* O9 has been proposed to be the result of recombination between the *wb** gene cluster *E. coli* O9 and that of *Klebsiella* O3. (Sugiyama et al, 1998)

2.3. LPS biosynthesis

Biogenesis of LPS is a very complex process involving many enzymes or enzyme complexes (Woodward et al, 2010), governed by more than forty genes (Valvano, 2003). The various steps occur at the plasma membrane and are followed by the translocation of LPS molecules to the bacterial cell surface (Valvano, 2003). Despite that structurally LPS is generally divided into three structural parts it is synthesized as two separate moieties: lipid A-core as one and OPS as another. The genes governing the synthesis of lipid A and substrates for the assembly of the IC components are scattered throughout the chromosome. In contrast, the genes encoding functions for OC and OPS are clustered (Valvano, 2003).

2.3.1. Biosynthesis of lipid A-core moiety

All steps in lipid A biosynthesis of *E. coli* have been well characterized and it is also known that the core is assembled in a stepwise manner onto lipid A (Raetz & Whitfield, 2002, Raetz et al, 2007). Lipid A biosynthesis starts by connecting one by one the two fatty acid chains to UDP-sugars (two 3OH-C14-ACP [acyl carrier protein] and UDP-GlcNAc). Then the two UDP-fatty acid sugar units are linked together and phosphorylated to form the precursor of lipid A called lipid IVa. Lipid IVa is decorated with core carbohydrates (Kdo) and more fatty acids to form Kdo2-lipid A. This part is the conserved part of the lipid A biosynthesis. The evidence to substantiate this includes results from elucidation of different lipid A structures, and the data generated from nucleotide sequencing. This information suggests that the lipid A biosynthesis follows a conserved general outline and especially the enzymes (LpxA and LpxC) catalyzing the first steps are the most conserved among bacteria. These intracellular enzymes are, present in virtually all Gram-negative bacteria and expressed constitutively (Raetz & Whitfield, 2002, Raetz et al, 2007).

Kdo2-lipidA serves as the acceptor on which the core oligosaccharide is assembled via sequential transfer of the glycosyl moieties from NDP-sugars. When OPS biosynthesis is completed (OPS biosynthesis will be discussed later), this proximal moiety is then linked to lipid A-core on the periplasmic face of the IM.

While the constitutive lipid A pathway discussed above is conserved the systems for lipid A modification are diverse and are usually not required for growth (Raetz et al, 2007, Raetz et al, 2009). Lipid A-modification enzymes are mostly extracytoplasmic and vary from organism to organism. They are often induced or repressed by growth conditions as discussed above (Raetz et al, 2007). Removal and modification of the lipid A phosphate groups occurs on the outer surface of the IM, whereas removal of acyl chains generally occurs in the OM. (Raetz et al, 2007, Raetz et al, 2009)

2.3.2. Transport of the LPS to the OM

LPS like all the other components of the OM are not synthesized *in situ,* so they must be transported to the OM from their site of synthesis. LPS is biosynthesized at the IM and until recently, little was known about its transport across the aqueous periplasmic compartment and assembly to the OM (Bos et al, 2004, Wang & Quinn, 2010). Now several proteins (LptABCDFG) have been reported to function in this process (Bos et al, 2004, Ruiz et al, 2008, Sperandeo et al, 2007, Sperandeo et al, 2008, Wu et al, 2006) but the molecular mechanism of the LPS transport across the periplasm is not well understood (Sperandeo et al, 2008). It is also unclear whether there are other components in the LPS transport machinery (Sperandeo et al, 2008). In addition, it will be challenging to characterize the molecular role of each protein and to understand the interactions of the proteins with each other (Sperandeo et al, 2008).

2.3.3. OPS biosynthesis

O polysaccharide is synthesized on an undecaprenyl phosphate (Und-P) carrier and then ligated to the lipid A-core structure. From a mechanistic standpoint, the biogenesis can be subdivided into four stages (i) the initiation reaction, (ii) the elongation/tanslocation/ polymerization of O-ag repeating subunits, (iii) the ligation to the lipid A–core and (iv) the recycling of the Und-P carrier (Valvano, 2003). Three main models for the OPS biosynthesis have been proposed (Raetz & Whitfield, 2002, Woodward et al, 2010) and these include the *wzy*-dependent pathway (also called heteropolymeric), the synthase-dependent pathway and the ABC transporter–dependent pathway. The ABC-transporter dependent pathway is also called homopolymeric even though in some cases the polymer does not contain a strictly homopolymeric structure but is composed of apparent di- or trisaccharides (Stenutz et al, 2006). In the two latter cases the entire O polysaccharide is synthesized at the cytoplasmic side of the IM. In ABC transporter–dependent pathway flippase consists of two different polypeptides (Wzm and Wzt), while in the synthase-dependent pathway, the synthase (WbbT) is the key enzyme. As the *wzy*-dependent pathway accounts for the synthesis of most heteropolymeric polysaccharides (including the YeO3 OC and YeO8 OPS) it is thus discussed further.

2.4. Biosynthesis of OPS by *wzy*-dependent pathway

In the *wzy*-dependent pathway the biosynthesis of the OPS is initiated on the cytoplasmic face of the IM by the addition of a sugar phosphate to Und-P to form monosaccharide-P-P-Und (Figure 2A). To this precursor the oligosaccharide repeating unit is assembled serially by different GTases which specifically transfer their target sugars from the donor NDP-sugars (Figure 2A). The newly constructed repeating unit is then translocated by a flippase (Wzx) to the periplasmic side of the IM (Figure 2B) where a polymerase (Wzy) catalyses their polymerization. The chain length regulator (Wzz) is postulated to enable the generation of strain-specific polysaccharide chain lengths (Raetz & Whitfield, 2002, Woodward et al, 2010) (Figure 2C). Finally the OPS is ligated to lipid A core (Figure 2D). In fact, the gene that encode Wzx, the putative flippase, is always present in the gene clusters containing the *wzy* gene and therefore Valvano proposed that it might be more appropriate to refer to this pathway as Wzy/Wzx-dependent pathway (Valvano, 2003). At present still both terms, *wzy*- and Wzy/Wzx –dependent pathway, are used.

2.4.1. Biosynthesis of NDP-sugars

Generally, the monosaccharides present in LPS are originating from corresponding NDP-sugar and there are specific enzymes catalyzing the formation of these NDP-sugars.

NDP-sugars are glycosyl esters of nucleoside diphosphates. Together with glycosyl esters of nucleoside monophosphates (NMP) they are called sugar nucleotides or glycosyl nucleotides. Sugar nucleotides can be distinguished into two classes, (i) primary and (ii) secondary glycosyl nucleotides. Primary glycosyl nucleotides are produced as a result of the interaction of nucleoside 5´-triphosphates with glycosyl phosphates or monosaccharides by nucleotidyltransferases (EC 2.7.7 group). Secondary glycosyl nucleotides are formed from the primary glycosyl nucleotides through enzymatic reactions leading to modification of their monosaccharide fragments. Most of the uncommon sugars present in LPS structures are result of transformations of the glycosyl group in glycosyl nucleotides (Shibaev, 1986). As this review concentrates on LPS only the formation of secondary NDP-sugars is discussed further.

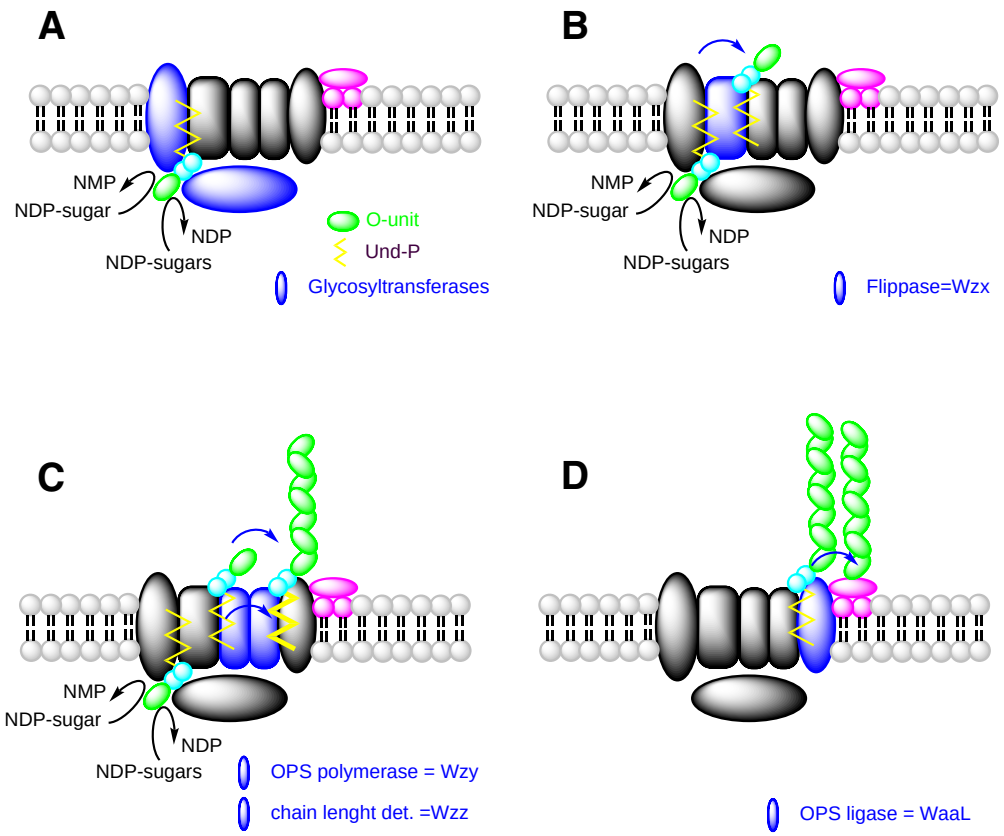

Figure 2. Biosynthesis of heteropolymeric OPS. A) O-unit synthesis B) O-unit flipping C) polymerization D) ligation. Modified from (Whitfield, 1995).

2.4.1.1. NDP-sugar synthesizing enzymes

For the bacteria to be able to synthesize LPS with different monosaccharide units it needs to have the correct kind of NDP-sugars available as building blocks. Some NDP-sugars are directly available as they are needed for other metabolic purposes of the bacteria but others are synthesized specifically for LPS biosynthesis and the genes for those enzymes are often localized to the gene clusters directing LPS biosynthesis.

There are several different reactions catalyzed by the group of enzymes here called as the NDP-sugar synthesizing enzymes. According to Thibodeaux *et al.* (Thibodeaux et al, 2007) the immediate biosynthetic precursor of most deoxysugars is an NDP-4-keto-6-deoxy-D-hexose (Figure 3) as most of the unusual sugars are 6-deoxy-hexoses. From this common intermediate, different combinations of oxidation, reduction, deoxygenation, epimerization, isomerization, group-transfer and rearrangement reactions yield the vast array of diverse NDP-sugar structures (Thibodeaux et al, 2007).

Many sugar biosynthetic enzymes have been studied but those involved in the deoxygenation are among the most mechanistically diverse and intriguing ones and thus

provided here as examples. The specific C-O bond cleavage strategy used by each type of deoxygenating enzyme is determined by the location of scissile C-O bond relative to an essential 4-keto group on the sugar ring. For instance, C6 deoxygenation is catalysed by an NDP-hexose-4,6-dehydratase also known as an oxidoreductase, a member of short-chain dehydrogenase/reductase (SDR) family, and C2 deoxygenation by TylX3 (from *Streptomyces fradiae)*. The C3 deoxygenation is more complex because the scissile C-O bond is adjacent to the 4-keto group and the reaction is catalyzed by combined action of two enzymes a 3-dehydrase (E_1) and a reductase (E_3). The mechanistic details for C4 deoxygenation are proposed but details remain obscure (Thibodeaux et al, 2007).

Many enzymes involved in the unusual sugar biosynthesis are members of the above mentioned SDR family. Enzymes of this family use a conserved protein fold and catalysis of these enzymes is initiated by oxidation of one of the sugar hydroxyl groups by the *nicotinamide adenine dinucleotide phosphate* $NAD(P)^+$ prosthetic group. SDR family enzymes catalyzing various epimerization reactions (C4, C2 and C6 epimerization by UDP-D-galactose-4-epimerase, cytosine diphosphate (CDP)-D-tyvelose-2-epimerase and adenine diphosphate (ADP)-β-L-*glysero*-D-*manno*-heptose-6-epimerase, respectively) have been characterized. Also the reaction mechanisms for SDR family enzymes catalyzing addition of sulphite group by UDP-α-D-sulphoquinovose synthase (SQD1), C6 decarboxylation by UDP-glucuronic acid decarboxylase, and C5 epimerization followed by C3 epimerization by guanine diphosphate (GDP)-D-mannose-3,5-epimerase/reductase, are known. These examples illustrate the versatility of this class of sugar biosynthetic enzymes. (Thibodeaux et al, 2007)

The biosynthesis of a particular NDP-sugar is often a cascade of several enzymatic reactions and thus some examples of the biosynthetic pathways for the LPS sugars are shown in Figure 4.

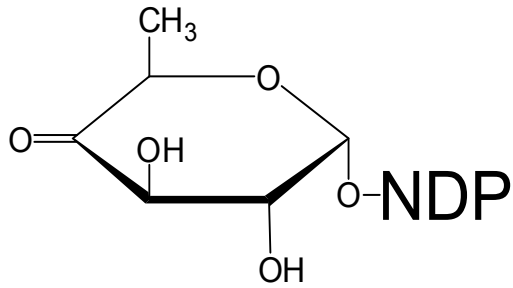

Figure 3. The structure of NDP-4-keto-6-deoxy-D-hexose, the immediate biosynthetic precursor of most deoxysugars.

2.4.2. Modifications of functional groups in monosaccharide units

As mentioned above, in addition to the various monosaccharide constituents many bacterial polysaccharides contain noncarbohydrate components. There are examples

where all the monosaccharide residues of a certain type in the polymer are modified and it seems reasonable to suggest that in these examples modifications take place prior to the formation of the polymeric chains (Shibaev, 1986). In those cases the modifications would be catalyzed by the NDP-sugar synthesizing enzymes.

The opposite situation, when not all of the monosaccharide residues of the polymer are modified (non-stoichiometric substitution), also exists (Shibaev, 1986). This results in the so-called masked regularity of the OPS, and the modifications are thought to take place after or in course of polymerization. (Knirel & Kochetkov, 1994, Knirel, 2010, Shibaev, 1986) In some OPS only the non-reducing end of the polymeric chain is modified, *e. g.* in *Klebsiella* O5 3-*O*-methyl-D-mannose is present in the non-reducing end of the polymeric mannose chain (Knirel & Kochetkov, 1994, Shibaev, 1986). In *E.coli* the methylation of the nonreducing terminal monosaccharide group of polymannose O-antigens is a signal for cessation of the chain elongation (Clarke et al, 2004).

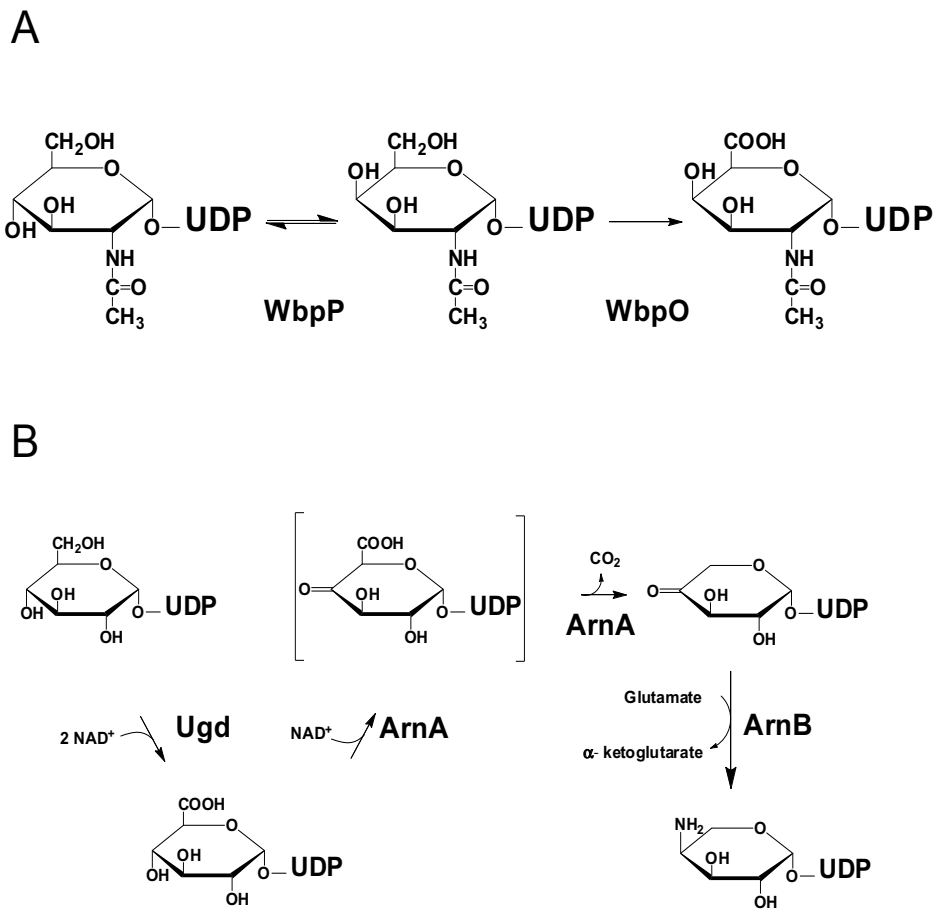

Figure 4. A) Biosynthesis of UDP-*N*-acetyl-D-galactosaminuronic acid (UDP-GalNAcA) in *Pseudomonas aeruginosa* serotype O6 (Creuzenet et al, 2000, Zhao et al, 2000).
B) Proposed pathway for UDP-L-Ara4N biosynthesis in *E. coli* (Raetz & Whitfield, 2002).

2.4.3. Glycosyltransferases (GTases)

The formation of glycosidic linkages is catalyzed by GTases. By definition, GTases are enzymes that utilize an activated donor sugar substrate that contains a phosphate-leaving group. Donor sugar substrates are usually activated in the form of NDP-sugars (but NMP-sugars, lipid phosphates and unsubstituted phosphate are also used). The nucleotide sugar dependent GTases are often called Leloir enzymes. In GTase reactions the acceptor substrates are most commonly other sugars but they can also be lipid, protein, nucleic acid, antibiotic or another small molecule. In addition although glycosyl transfer most frequently occurs to the nucleophilic oxygen of a hydroxyl substituent of the acceptor it can also occur to nitrogen, sulfur and carbon (Lairson et al, 2008).

In the case of GTase reactions of LPS biosynthesis, glycosyl nucleotides (NDP-sugars) are the usual donors of the reaction and the transfer occurs to the nucleophilic oxygen of the acceptor, which is the carbohydrate-containing polymer (the intermediate of the Lipid A-core biosynthesis or the mono- or oligosaccharide on Und-P lipid carrier in the biosynthesis of the OPS) (Raetz & Whitfield, 2002). Inter-monomeric linkages in bacterial polysaccharide chains are very diverse and more than 400 different GTases are necessary in order to ensure their synthesis (Shibaev, 1986). Bacterial GTases involved in the synthesis of exopolysaccharides have an abundance of positively-charged amino acids and they account at least in part for the ionic interactions of these proteins with the plasma membrane (Valvano, 2003).

GTases have been classified to sequence-based families by Campbell (Campbell et al, 1997) and the classification has been further developed by Coutinho (Coutinho et al, 2003). The continuously updated information is available in the Carbohydrate-Active EnZymes database (Cazy) (http://www.cazy.org/).

In the glycosylation reaction the stereochemistry at the C1 position of the donor sugar (UDP-sugar) can remain or change. According to that the GTases are either retaining or inverting, respectively. However, a reliable prediction of the catalytic mechanism (inverting or retaining) is not always possible based on sequence comparison alone (Breton et al, 2006). The mechanism of inverting GTases is that of a straightforward S_N2-like reaction facilitated by an enzymatic base catalyst and by Lewis acid activation of the departing (substituted) phosphate leaving group. The mechanism of retaining GTases remains less clear (Lairson et al, 2008).

Accordingly to the solved X-ray structures GTase folds have been observed to consist primarily of $\alpha/\beta/\alpha$ sandwiches. In addition, GTases seem to mainly fall in two structural superfamilies: GT-A and GT-B. Inverting and retaining GTases are found in both superfamilies.

GT-A family GTases seem to have two characteristic regions. The first region (100-120 N-terminal residues) corresponds to the Rossmann-type nucleotide binding domain ($\alpha/\beta/\alpha$ sandwich) and it is terminated by a general feature of the GT-A family: DxD motif.

The DxD motif has been shown to interact with the phosphate groups of the nucleotide donor through the coordination of a divalent cation, typically Mn^{2+}. The C-terminal portion of the GT-A GTase is highly variable, mostly dedicated to the recognition of the acceptor, but a common β–α–α structural motif forming a part of the active site is seen. In inverting enzymes, the presumed catalytic base has been proposed in this region. The ratio of loops to secondary elements is high in GTases and the flexible loops appear to be important for the substrate binding (Breton et al, 2006). It is also worth noting that not all enzynmes that possess GT-A fold are GTases, and not all GT-A fold GTases possess the DxD signature motif. Furthermore, many proteins possess a DxD signature motif in their sequences but these are not necessarily GTases.

The specificity of GTases can vary. Generally one GTase catalyses one reaction and for every glycosidic linkage present in OPS unit or core there is a specific GTase available. Sometimes the specificity is really surprising like in *E. coli* core types R1 and R4 where WaaV and WaaX add different β-linked sugars (Glc and Gal, respectively) to the same monosaccharide unit of the otherwise identical core. (Heinrichs et al, 1998a, Raetz & Whitfield, 2002)

The enormous sequence data available nowadays has allowed one to predict functions for uncharacterized enzymes according to their sequence similarities, but especially in the case of GTases to definitively prove their specificities experimental evidence is needed. For instance, two urdamycin A GTases (UrdGT1b and UrdGT1c) which have 91 % sequence identity yet show different specificities for both sugar and acceptor substrates. (Thibodeaux et al, 2007)

In general, GTases are considered monofunctional and highly specific but exceptions do exist. Concerning LPS biosynthesis, Kdo transferases is an example, and mono-, bi- tri- and even tetrafunctional Kdo transferases have been characterized. The guideline is that a certain multifunctional Kdo transferase is proposed to recognize different acceptors in consecutive reactions (Gronow et al, 2010).

This multifunctionality should be distinguished from polymorphism. In polymorphism genes from different strains have same name derived from their sequence similarity but they can have different specificities as already single amino acid difference can change the specificity. As an example the LspA enzymes from different *Haemophilus influenzae* strains have been shown to direct addition of Glc or Gal to the terminal inner core heptose via alternative linkages (Deadman et al, 2006).

2.4.4. Primase

When O-ag unit is synthesized on Und-P lipid carrier in the initiation reaction, the first carbohydrate is transferred together with one P to Und-P from NDP-sugar to form sugar-P-P-Und. UMP is released in the reaction and the catalyzing enzyme is sometimes called a primase. The linkage formed is not identical to the glycosidic linkages used to build up the rest of the repeat unit and the enzyme catalyzing the initiation reaction belongs to

prenyl sugar transferases (Lehrer et al, 2007). There are two different classes of integral membrane proteins catalyzing the initiation reactions (O-ag initiation as being one kind of initiation reaction): polyprenyl-phosphate hexose-1-phosphate transferases (PHPT) and polyprenyl-phosphate *N*-acetylhexosamine-1-phosphate transferases (PNPT) (Valvano, 2003).

PHPT uses NDP-hexoses as substrates and WbaP serves as a prototype member of this family. PNPTs on the other hand use NDP-hexosamines as substrates and WecA provides a prototype. PNPT superfamily includes both eukaryotic and prokaryotic members. Bacterial members of this superfamily are implicated in the biosynthesis of different cell envelope polymers (f.ex. O-ag, ECA, peptidoglycan or teichoic acids) and share a common membrane bound acceptor substrate C_{55}-P but as mentioned above utilize different UDP-*N*-acetyl-hexosamine substrates (Bouhss et al, 2008).

As a prototype member, the WecA protein is discussed here in more detail. The *wecA* gene is located within ECA gene cluster and WecA serves to initiate both O-antigen and ECA synthesis (Rick et al, 1998, Valvano, 2003). It has been shown that WecA is necessary for both Wzx/Wzy-dependent and Wzx/Wzy-independent O-subunit pathways (*e. g.* in *E. coli* K-12 and *E. coli* O8 and O9, respectively) (Valvano, 2003). By analyzing fluorescence labeled WecA it has been shown to have punctate distribution around the cell perimeter. This supports the hypothesis that there are multiprotein complexes for capsular polysaccharides serving as molecular scaffold across the periplasm (Lehrer et al, 2007). As most heteropolymeric O-antigen repeating units have either GlcNAc or GalNAc at the reducing terminus it has been assumed that biosynthesis of the lipid intermediates is initiated by the transfer of either GlcNAc-P or GalNac-P (from UDP-GlcNac or UDP-GalNAc, respectively) to Und-P by WecA. The conclusion that in addition to UDP-GlcNAc, WecA would utilize also UDP-GalNAc has had no experimental evidence. In fact Rush *et al.* have recently shown that at least in *E. coli* (strain O157) WecA does not have enzymatic activity for UDP-GalNAc but that the GalNAc-P-P-Und would be synthesized from GlcNAc-P-P-Und by an epimerase (Z3206). (Rush et al, 2010)

2.4.5. O-unit flippases (Wzx)

As the O-unit is synthesized on the lipid carrier on the cytoplasmic side of the IM it must be translocated to the periplasmic side of the membrane before it can be attached to the lipid-A core. This translocation is proposed to be catalyzed by a flippase protein called Wzx.

Proteins mediating the excretion of complex carbohydrates into or across the bacterial cell envelope are classified to two principal families; polysaccharide-specific transport (PST) systems and ATP-binding cassette-2 (ABC-2) systems. Wzx proteins are predicted to have 12 transmembrane helices and they are classified within a subfamily of PST systems, PST(2) (Paulsen et al, 1997). Even though Wzx proteins reveal very little sequence conservation they do share similar motifs (Samuel & Reeves, 2003)

and different Wzx proteins have been shown to complement each other to some extent (Feldman et al, 1999).

Wzx proteins from several O-ag systems appear to have specificity for the first sugar attached to und-P, regardless of the chemical structure of the remainder of the subunit (Marolda et al, 2004). Recognition of a complete Und-PP bound O-ag subunit is not required for the translocation, in fact a single sugar (GlcNAc) can be incorporated to LPS (of *E. coli* K-12) (Feldman et al, 1999).

When the different complementation abilities of Wzx proteins from different O-ag systems were further characterized it was proposed that Wzx proteins would have sub-families, some interact with O-units that are initiated with *N*-acetylhexosamines and others might better interact with O-units initiated with hexoses (Marolda et al, 2004).

2.4.6. OPS polymerases (Wzy)

After lipid carrier-linked O-repeat units (Und-PP-O unit) are translocated to the periplasmic side of the IM they are polymerized to Und-PP linked OPS. This Wzy-dependent polymerization of the OPS occurs at the reducing end of the nascent chain being formed, meaning that the OPS on the Und-PP carrier is transferred to the most recently synthesized Und-PP-oligosaccharide (Robbins et al, 1967, Stenutz et al, 2006). This reaction has been proposed to be catalyzed by an integral membrane protein called OPS polymerase Wzy.

Wzy proteins appear to possess a relative large periplasmic loop and exhibit little primary sequence homology between each other (Daniels et al, 1998, DebRoy et al, 2004, Guo et al, 2008). The configuration of the formed glycosidic linkage connecting the O-units can be either α- or β, and in *P. aerigunosa* the different Wzy proteins are referred as Wzy_{α} and Wzy_{β}, respectively (Kaluzny et al, 2007). Unlike typical GTase catalyzed reactions where P-P-lipid moiety is part of the acceptor structure in the Wzy catalyzed polymerization the P-P-lipid moiety also serves as part of the donor and it is directly involved in the bond breaking in the reaction. Recognition for Wzy toward the lipid structure of the sugar donor has been illustrated and a defined mode of it has been suggested (Woodward et al, 2010).

Despite extensive genetic studies, only recently Woodward *et al.* were able to demonstrate an *in vitro* reconstitution of *E. coli* O-polysaccharide biosynthesis using chemically defined substrates and purified biosynthetic enzymes and thus provided first direct biochemical evidence that Wzy is both necessary and sufficient to induce polymerization that is to say that Wzy is the polymerase. (Woodward et al, 2010)

2.4.7. O chain length determinant (Wzz)

The lengths of the polysaccharides on the cell surface of bacteria are usually not random but distribute around some mean value, termed modal length. A large family of proteins

called polysaccharide co-polymerases (PCPs) regulate polysaccharide lengths (Morona et al, 2000, Morona et al, 2009). Common features of all PCPs are the presence of N-terminal and C-terminal transmembrane helices separates by a large hydrophilic region with a predicted coiled-coil region (Marolda et al, 2008, Morona et al, 2000, Tocilj et al, 2008). Wzz proteins belong to the PCP 1 group and Wzz that operate on O-ag are distinguished from the Wzz proteins that regulate the ECA lengths to subgroups PCP 1a and PCP 1b, respectively (Barr et al, 1999). In PCP 1 Wzz proteins the hydrophilic periplasmic region characteristically has 240-280 amino acids (Burrows et al, 1997, Franco et al, 1996, Morona et al, 1995, Tocilj et al, 2008) and despite low sequence identity it has been proposed that all Class 1 PCPs have similar protomer structure (Tocilj et al, 2008). The probability of the coiled coil in Wzz protein seems to correlate with the chain length of the synthesized O-ag: when the probability of coiled coil is high then longer O-antigen chains are synthesized. (Guo et al, 2008, Paulsen et al, 1997).

PCPs are known to associate to oligomers to accomplish the task of controlling oligosaccharide lengths (Daniels & Morona, 1999, Daniels et al, 2002) but the mechanism by which Wzz proteins impact the polymerization process is presently unknown (Morona et al, 2009). Several models have been proposed (Bastin et al, 1993, Morona et al, 1995), and it has been shown that Wzz operates on the lipid carrier linked polymer prior the ligation step by WaaL (Daniels et al, 2002).

Recently after analyzing the full-length Wzz proteins Larue *et al.* proposed that Wzz could modulate the O-chain length by regulating the flow of repeat units into the polymerization reaction but detailed understanding of the molecular interactions between Wzz and other component in the system is needed to fully solve the problem (Larue et al, 2009, Woodward et al, 2010).

Biochemical evidence for the *E. coli* Wzz to act as chain length regulator was provided recently (Woodward et al, 2010). And although Wzz proteins are not essential for O-ag biosynthesis, their ability to control the O-ag length is crucial for bacterial virulence (Bengoechea et al, 2004, Hong & Payne, 1997, Murray et al, 2003, Murray et al, 2006, Tocilj et al, 2008).

2.4.8. LPS ligases

Newly-synthesized OPS is attached to lipid A core on the periplasmic side of the IM by an enzyme called OPS ligase. According to Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB) the E.C. 6 subclass defines ligases as enzymes that catalyze the joining of two molecules with concomitant hydrolysis of the diphosphate bond in ATP or a similar triphosphate. The ligases forming carbon-oxygen bonds are members of the E.C. 6.1. subclass and the ligases forming phosphoric ester bonds are members of the E.C. 6.4. According to these definition criteria the ligases adding OPS, OC and ECA to Lipid A (–core) would be members of subclass E.C.6.1 but

when ECA is ligated so that phosphoric ester bond is formed the ligase would belong to E.C. 6.4.

Just like glycosyltransfereses LPS ligase enzymes form also glycosidic linkages. The difference is the donor. For ligases the donor is an undecaprenylpyrophosphate –linked oligo- or polysaccharide as for GTase the donor is a nucleotide sugar. Consequently ligases share no similarity with GTases (Heinrichs et al, 1998b).

The ligase proteins are called WaaL and they are integral membrane proteins with eight or more membrane-spanning domains. The amino acid sequences of different WaaL proteins exhibit low level of similarity. (Abeyrathne et al, 2005, Heinrichs et al, 1998b). WaaL are currently the only proteins to be known to be involved in the ligation of O-ag from Und-PP-carrier to the lipid A core. Little is known concerning the process or the requirement for additional cellular factors. WaaL proteins seem to require a specific lipid A -core oligosaccharide acceptor structure for ligase activity and from the donor they rather recognize the Und-P carrier than the saccharide attached to it (Guo et al, 2008, Heinrichs et al, 1998b). However specificity of the ligation reaction for a particular core structure is not dependent on the WaaL protein alone (Kaniuk et al, 2004).

Until quite recently when Abeyrathne *et al.* (Abeyrathne & Lam, 2007) were able to provide biochemical evidence of the ligation activity of WaaL in *P. aerigunosa* only genetic data had implicated that *waaL* is involved as an O-antigen ligase in the LPS assembly. In the *in vitro* ligase assay the O-antigen ligation was shown to be a energy-dependent process and that a key amino acid residue H303 in the large periplasmic loop region is important for ligase activity (Abeyrathne & Lam, 2007). In a separate investigation of WaaL in *E. coli*, Perez *et al.* (2008) stated that the conserved residues of the characteristic periplasmic loop was unlikely to participate in the recognition of core OS terminal sugars but would participate in the recognition of the Und-PP moiety of the Udd-PP linked O antigen. These authors suggested that WaaL likely participates in the chemical reaction or reactions required for the release of O antigen from the Und-PP carrier (Perez et al, 2008). The controversy between the two studies could not be resolved until there are more biochemical data to define the energy requirement and the mechanism of the O-antigen ligation process.

2.4.9. Possibility of multi protein complexes

Wzx, Wzz and Wzy presumably form a multi protein complex (Marolda et al, 2006) and the recently reported *in vitro* reconstitution system by Woodward *et al.* (Woodward et al, 2010) further confirmed such among Wzy, Wzz and Und-P linked polymers. It has also been proposed that WaaL is another component complexed with Wzx, Wzy and Wzz (Feldman et al, 1999, Kaniuk et al, 2004), however there is no conclusive evidence of protein-protein interactions involving WaaL (Kaniuk et al, 2004). Some kind of interplay involving at least some of the above mentioned proteins seems to be obvious e.g. in *Shigella flexneri* 2a that has two preferred O-ag chain lengths the expression level of

Wzy has been shown to be critical for the O-antigen modal distribution (Carter et al, 2009) but the detailed picture of the interaction mechanisms of the proteins involved in the O-ag assembly from the Und-PP- carrier on the cytoplasmic side of the IM to the lipid A core on the periplasmic side of the IM remains to be solved.

2.4.10. The recycling of Und-PP

The availability of Und-P is a limiting factor in the biosynthesis of OPS, since this lipid carrier is made in very small amounts and it is also required for biosynthesis of other carbohydrate polymers (Valvano, 2003, Valvano, 2008). Antibiotics like bacitracin and colicin M have inhibitory effect on OPS and peptidoglycan biosynthesis by preventing the recycling of Und-PP. The Wzy/Wzx dependent pathway is more severely affected by these antibiotics as it involves one molecule of Und-PP per every O repeating unit. In contrast ABC transport and synthase-dependent pathways theoretically require only one Und-P molecule per every polymer chain, thus are less affected by the above two antibiotics. (Valvano, 2003)

For a long time the knowledge about the enzyme or enzymes which carry out the Und-PP dephosphorylation on the periplasmic side of the IM were not known; however, in 2008 Tou´ze (Touze et al, 2008) discovered that purified LpxT catalyses the transfer of a phosphate group from Und-PP donor to the 1-position of the lipid A to form lipid A 1-diphosphate in a reaction that occurs in the periplasmic side of the membrane (Tatar et al, 2007). These observations provided a novel and unexpected link between periplasmic lipid A modifications and the Und-PP recycling pathway; however, the functional significance of this process remains to be elucidated (Valvano, 2008). The mechanisms and molecules involved in the translocation of the periplasmic-generated Und-P to the cytosolic side of the membrane where it can serve as a glycan acceptor again remains to be elucidated (Valvano, 2008).

2.5. LPS as biological compound, special focus on YeO3 OC and YeO8 O-ag

The location of LPS at the cell surface places it at the interface between the bacterium and its environment. The bottom-up increasing level of diversity of LPS can be explained with the selective pressures from host responses, bacteriophages and environmental stresses exerted onto the outer most parts of the bacterial surface. The important roles of lipid A and IC in the cell envelope properties are more closely linked to the maintenance of OM integrity, while the primary role of O polysaccharide appears to be affording a protective shield/shell to the bacteria. This phenomenon applies also to YeO3 and YeO8 LPS.

2.5.1. Lipopolysaccharide of YeO8

The structure of YeO8 LPS is almost completely elucidated. The compositions of Lipid A and the core region have been reported by Oertelt *et al.* (Oertelt et al, 2001, Perez-Gutierrez et al, 2010). The lipid A part consists mainly of a tetra-acylated and biphosphorylated glucosamine (GlcN) disaccharide, harboring two C14:0(3OH) residues on the reducing GlcN and one amide-linked acyloxyacyl (C14:0[3-O-C14:0]) residue on the nonreducing GlcN. Recently Perez-Gutierrez *et al.* have reported that in fact the acylation of YeO8 lipid A is temperature dependent, at 37 ºC tetra-acylated and at 21 ºC hexa-acylated lipid A predominates (Perez-Gutierrez et al, 2010).

The carbohydrate backbone of the major portion of the core region of *Y. e.*O8 LPS is identical to that of LPS from *Y. e.* O:3 (see below). However, in YeO8 the terminal D,D-Hep is not nonstoichiometrically substituted with phosphate as it is in the YeO3 IC. Also, in the YeO8 core, a branching β-linked GlcNAc was found to be linked to O3 of the middle L,D-Hep residue. (Oertelt et al, 2001)

In YeO8 the OPS is heteropolymer containing branched pentasaccharide repeating units of GalNAc, Gal, mannose (Man), fucose (Fuc) and 6d-gulose (Gul) (Tomshich et al, 1987) but the linkage between the OPS and the lipid A-core is at present unknown.

The gene cluster directing O-ag biosynthesis of YeO8 has been cloned, sequenced and characterized (Bengoechea & Skurnik, 2000, Bengoechea et al, 2002, Zhang et al, 1996, Zhang et al, 1997), but experimental evidence to conclusively prove the proposed biosynthesis pathway is still missing. Also the biosynthesis of lipid A-core region remains to be experimentally proven. However taken into account the conserved structure of lipid A in *Enterobacteriaceae* the biosynthesis of this innermost moiety of LPS can be assumed to follow the general guidelines established in other bacteria and described above.

2.5.2. Lipopolysaccharide of YeO3

The structure of YeO3 LPS is partly solved. The structure of Lipid A is identical to the one of YeO8 (Aussel et al, 2000, Oertelt et al, 2001). The IC is a branched heptasaccharide with D,D-Hep(1,7)-L,D-Hep(1,7)-L,DHep(1,3)-L,D-Hep(1,5)Kdo backbone where two Glc residues are β-linked to C4 and C2 of the innermost and second innermost L,D-Hep residues, respectively. In addition the D,D-Hep is partially phosphorylated at C7 (Radziejewska-Lebrecht et al, 1994). The YeO3 produces a homopolymeric OPS (6-deoxyaltrose as the monosaccharide unit) and a hexasaccharide branch known as OC and reported to contain two Glc, one Gal, one *N*-acetylfucosamine (FucNac) and two *N*-acetylgalactosamines (GalNAc) (Radziejewska-Lebrecht et al, 1998). A special characteristic in *Ye* O:3 is that both OPS and OC are linked to lipid A core, OC is linked to the C3 of the second innermost L,D-heptose (Radziejewska-Lebrecht et al, 1998) but the position and exact linkage connecting OPS and IC is unknown. There is also some uncertainty concerning the first monosaccharide unit of the OC that previously

was suggested to be FucNAc (Radziejewska-Lebrecht et al, 1998). YeO3 core and OPS structures are shown in Figure 1. Both the YeO3 OPS and OC gene clusters have been cloned, sequenced and characterized (al-Hendy et al, 1991b, Skurnik et al, 1995), but experimental evidence for the functions of the gene products have not been presented.

2.5.3. Virulence factor

LPS has for long been implicated as a virulence factor and in fact the early interest in the chemistry, biosynthesis and genetics of LPS was stimulated by their potential application in vaccine development.

Both YeO3 OC and YeO8 OPS are shown to be virulence factors (Najdenski et al, 2003, Skurnik et al, 1999, Zhang et al, 1997). YeO3 OC plays a role in virulence early after entry in the host; most likely in resistance against bactericidal peptides of macrophages and polymorphonuclear phagocytes (Skurnik et al, 1999). It has been proposed that YeO3 OC operon would be a relic of an ancestral heteropolymeric OPS operon and it is possible that the critical role of OC might explain why it has not been completely replaced by the homopolymeric OPS gene cluster (Skurnik et al, 1999).

2.5.4. Endotoxin

Lipid A part of the LPS is called an endotoxin. Most type of lipid-A molecules are detected at pico molar levels by ancient receptor of the innate immune system (most significantly by Toll-like receptor 4, TLR4). Under favourable situations activation of TLR4 by lipid A leads to clearing local infections. However, in the case of severe sepsis, when the mediators are overproduced, the outcome can be serious (Raetz & Whitfield, 2002). In human cells, the two phosphate groups and acyloxyacyl moieties are especially needed for triggering the endotoxic response. Fot the first time the important role of endotoxin of YeO3 in infection was reported by Kanamori 1976 (Kanamori, 1976).

2.5.5. Receptor for bacteriophages and enterocoliticin

Different structural parts of LPS serve as receptors for different bacteriophages and bacteriocins. Enterocoliticin is a channel-forming bacteriocin produced by *Y. e.* 29930 (biogroup 1A; serogroup O:7,8) that uses the YeO3 OC as its receptor (Strauch et al, 2003); it kills enteropathogenic strains of *Y. e.* belonging to serogroups O:3, O:5,27 and O:9 (Strauch et al, 2001).

The most studied YeO3-specific phages are ϕR1-37 and ϕYeO3-12. ϕYeO3-12 is known to use the YeO3 OPS as its receptor (Pajunen et al, 2000) and the receptor for ϕR1-37 is YeO3 OC (Kiljunen et al, 2005). The minimum requirements and the possible need for involvement of some inner core structures to both the recognition of enterocoliticin and to ϕR1-37 are not known.

In addition to YeO3, other *Y. e.* serotypes (O:1, O:6, O:6,31, O:9, O:25,26,44, O:41,43 and O:50) have been reported to be sensitive to ϕR1-37 (Skurnik et al, 1995). Also most of the tested strains of the enterocoliticin sensitive serotype O:5,27 were ϕR1-37 sensitive. The enterocoliticin producing serotype O:7,8 is ϕR1-37 resistant (Skurnik et al, 1995).

2.5.6. Temperature regulation

Most of the OPS gene operons appear to be constitutively expressed and they are often preceded by a JUMPStart (Just Upstream of many Polysaccharide associated gene starts) /ops (operon polarity suppressor) sequence indicating that an anti-termination process is active. YeO3 appears to be an exception, since its OPS biosynthesis locus is transcriptionally regulated by temperature and regulation is mediated by a regulator encoded outside of the cluster (al-Hendy et al, 1991a, Raetz & Whitfield, 2002). Also the expression of YeO8 OPS is temperature regulated (Bengoechea et al, 2004). Temperature regulation has been proposed to be beneficial for virulence, the OPS expressed at room temperature is longer than the one expressed at 37 ºC (the body temperature of the host) (Skurnik & Toivanen, 1993). Contrary YeO3 OC biosynthesis genes are not reported to be thermoregulated (al-Hendy et al, 1991a, Skurnik & Toivanen, 1993).

3. AIMS OF THE PRESENT STUDY

The major goal of this study was to get a comprehensive view of the biosynthesis of YeO3 OC. As some of the results of the original publications are outside this frame I have excluded them from my PhD study. On the other hand some unpublished data and results from additional publications are included and discussed as they have had a major role in the success of this work.

The specific aims are listed below, the corresponding original publication/s are shown in parenthesis.

- Characterization of the UDP-*N*-acetylglucosamine 4-epimerase activity in YeO3 and YeO8 (I)
- Functional characterization of the WbcP protein (II)
- Characterization of GTases of the YeO3 OC gene cluster (III)
- Identification and characterization of the oligosaccharide ligase needed to ligate YeO3 OC to the IC (IV)
- Characterization of receptor properties of YeO3 OC (II, III)

4. MATERIALS AND METHODS

The detailed descriptions of the materials and methods used in this study are given in the original publications I-IV.

4.1. Bacterial strains and plasmids

Bacterial strains and plasmids used in this work are listed in Tables 3 and 4, respectively. Typically, *Y. e.* and *E. coli* were grown in liquid cultures in tryptic soy broth (TSB) or Luria-Bertani (LB) medium at 22 ºC to 25 ºC (room temperature, RT) or 37 ºC, respectively. Luria agar (LA) was used for solid cultures for both *Yersinia* and *E. coli* and when necessary cefsulodin-irgasan-novobiocin (CIN) agar was used as selective medium for *Yersinia*. When appropriate, antibiotics were added to the growth medium. The construction of the plasmids as well as the different bacterial strains is described in the original publications (I-IV).

Table 3. Bacterial strains used in this work.

Strain	Genotype	Reference
Yersinia enterocolitica		
6471/76 (=YeO3)	Wild type strain, patient isolate	(Skurnik, 1984)
6471/76-c (=YeO3-c)	Virulence plasmid cured derivative of 6471/76	(Skurnik, 1984)
YeO3-R1	Rough (OPS neg) derivative of YeO3-c	(al-Hendy et al, 1992)
YeO3-c-wbcN1	YeO3-c wbcN1 (the *wbcN* gene interrupted by the modified SalI site)	(Sirisena & Skurnik, 2003)
YeO3-c-wbcN1-R	Rough (OPS neg) derivative of YeO3-c-wbcN1	III[a]
YeO3-wbcO1	YeO3 wbcO1 (the *wbcO* gene interrupted by the modified NsiI site)	(Sirisena & Skurnik, 2003)
YeO3-wbcO1-R	Rough (OPS neg) derivative of YeO3-wbcO1	III
YeO3-c-wbcQ1	YeO3-c wbcQ1 (the *wbcQ* gene interrupted by the modified BglII site)	(Sirisena & Skurnik, 2003)
YeO3-c-wbcQ1-R	Rough (OPS neg) derivative of YeO3-c-wbcQ1	III
YeO3-c-OC	Δ(*wzx-wbcQ*), derivative of YeO3-c	(Biedzka-Sarek et al, 2005)
YeO3-c-OCR	Rough (OPS neg) derivative of YeO3-c-OC	(Biedzka-Sarek et al, 2005)
Ye75S		(Acker et al, 1981)
YeO3-c-trs22-R	YeO3-c Δ*wbcKL*	(Skurnik et al, 1999)
YeO3-c-trs24-R	YeO3-c Δ*wbcP*	(Skurnik et al, 1999)
YeO3-c-lig1727	$waaL_{os}$::pEPlig1727su, Clm^{Rb}	IV
YeO3-c-lig532	$waaL_{ps}$::pEPlig532su, Clm^{R}	IV
YeO3-c-lig777	$waaL_{eca}$::pSW25-lig777del, Spe^{R}	IV
YeO3-R1-lig1727	$waaL_{os}$::pEPlig1727su, Clm^{R}	IV
YeO3-R1-lig532	$waaL_{ps}$::pEPlig532su, Clm^{R}	IV
YeO3-c-lig1727-lig532	$waaL_{os}$::pEPlig1727su $waaL_{ps}$::pSW29-lig532del, Clm^{R}, Kan^{R}	IV

Strain	Genotype	Reference
YeO3-c-lig1727-lig777	*waaL*$_{os}$::pEPlig1727su *waaL*$_{eca}$::pSW25-lig777del, Clm^R, Spe^R	IV
YeO3-c-lig532-lig777	*waaL*$_{ps}$::pEPlig532su *waaL*$_{eca}$::pSW25-lig777del, Clm^R, Spe^R	IV
YeO3-c-lig1727-lig532-lig777	*waaL*$_{os}$::pEPlig1727su *waaL*$_{ps}$::pSW29-lig532del *waaL*$_{eca}$::pSW25-lig777del, Clm^R, Kan^R, Spe^R	IV
YeO3-c-lig532-SW-3	*waaL*$_{ps}$::pSW29-lig532del, Kan^R	IV
8081-R$^-$M$^+$ (YeO8)	R$^-$M$^+$ derivative of wild-type strain 8081; serotype O:8; pYV+	(Zhang et al, 1997)
8081-c R$^-$M$^+$ (YeO8c)	R$^-$M$^+$ derivative of 8081-c; serotype O:8; the pYV-cured derivative of 8081	(Zhang & Skurnik, 1994)
8081-R1	Spontaneous rough mutant strain	(Zhang et al, 1997)
Escherichia. coli		
C600	*thi thr leuB tonA lacY supE*	(Appleyard, 1954)
DH10B	F$^-$ *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC), ϕ80*lacZ*ΔM15 Δ*lac*X74, *deo*R, *recA*1 *endA*1 *ara*Δ139 Δ*(ara, leu)*7697 *galU, gal*K λ$^-$ *rpsL nupG* λ$^-$ *tonA*	Life Technologies
NovaBlue(DE3)	*endA1 hsdR17(r*$_{K12}^-$*m*$_{12}^+$*)supE44 thi*-1 *recA1 gyrA96 relA1 lac* [F′*proA*$^+$*B*$^+$*lacI*^{q}*Z*ΔM15::Tn*10*(TcR)] (DE3)	Novagen
JM109	F′ *traD*36 *proA*$^+$*B*$^+$ *lacI*^{q} Δ(*lacZ)*ΔM15/ Δ(*lac-proAB*) *glnV*44 *e*14$^-$ *gyrA*96 *recA*1 *relA*1 *endA*1 *thi hsdR*17	(Yanisch-Perron et al, 1985)
HB101/pRK2013	Conjugation helper strain, Kan^R	(Ditta et al, 1980)
AD9	Δ(*galK-bioD*)76 *relA1 rpsL125 thi-1 arg174;* Rc-type LPS	(Kikuchi et al, 1975)
DH5α	F$^-$Φ80*lacZ*Δ*M15 endA recA hsdR*($r_K^-m_K^-$) *supE thi gyrA relA1* Δ (*lacZYA-argF*)*U169*	(Woodcock et al, 1989)
BL 21 (DE3)	F$^-$ *ompT,hsdS*$_B$*(r*$_B^-$*m*$_B^-$*) gal dcm* (DE3)	Novagen
S17-1λ pir	*A-pir* lysogen of S17-1*, E. coli* strain for suiside vector delivery	(Wilson et al, 1995)
ω7249	B2163Λnic35, *E.coli* strain for suiside vector delivery, requirement for diaminopimelic acid 0.3mM, KanR	(Babic et al, 2008)
BW19610	DE3(*lac*)X74 Δ*uidA*::*pir*-116 *recA*1 Δ*phoA*532 Δ(*phnC?D-P*)33-30	(Metcalf et al, 1994)
W3110	*rph-1* IN(*rrnD-rrnE*), wt control for *waaL* mutant	(Bachmann, 1972)
CLM24	W3110, Δ *waaL*	(Feldman et al, 2005)

a) roman numerals refer to original publications

b) Clm, chloramphenicol; Spe, spectinomycin; Kan, kanamycin.

Table 4. Plasmids used in this work.

Plasmid	Genotype	Reference
pEEgne	YeO8 *gne* in pTM100, under the control of the tet gene promoter; Clm^{Ra}	I[b]
pEEgne-R	YeO8 *gne* in pTM100, opposite orientation to pEEgne; Clm^R	I
pEEgne-Y_1	*gne1*; Leu136Tyr; Clm^R	I
pEEgne-Y_2	*gne2*; Cys297Tyr; Clm^R	I
pEEgne-Y_1Y_2	*gne3*; Leu136Tyr and Cys297Tyr; Clm^R	I
pUC18-1	YeO8 *gne* and *wzz*, Amp^R	I
pRV28	YeO3 *gne* in pTM100, under the control of the Tet gene promoter; Clm^R	I
pETWbcP	full length *wbcP* of YeO3-c	II
pETWbcP-S212	YeO3-c WbcP Δ(1-211)	II
pETWbcP-E243	YeO3-c WbcP Δ(1-242)	II
pEPWbpVc	*P. aeruginosa* serotype O6 strain (ATCC 33354) *wbcV* in pTM100, under the control of the Tet gene promoter; Clm^R	II
pEPWbpVRe	*P. aeruginosa* serotype O6 strain (ATCC 33354) *wbcV* in pTM100, reverse orientation	II
pEPWbpK	*P. aeruginosa wbpK* from pET-28a-wbpK (Miller, 2006) in pTM100, under the control of the Tet gene promoter; Clm^R	II
pRV16NP	YeO3 OC gene cluster cloned in pTM100, Clm^R	(Kiljunen et al, 2005)
pRV16NPwbcK2	*wbcK-1* a non-polar frame shift mutation introduced by a 4 bp deletion at codon 142 overlapping the PvuI site of *wbcK*	III
pRV16NPwbcL1	*wbcL-1* a non-polar frame shift mutation introduced by a 4 bp deletion at codon 142 overlapping the SacI site of *wbcL*	III
pRV16NPwbcM1	*wbcM-1* a non-polar frame shift mutation introduced by a 4 bp deletion at codon 106 overlapping the KpnI site of *wbcM*	III
pRV16NPwbcQ1	*wbcQ-1* a non-polar frame shift mutation introduced by a 4 bp insertion at codon 219 overlapping the BglII site of *wbcO*	III
pRV16NPwzx	*wzx* a non-polar frame shift mutation	this study[c]
pYeO3wbcK	The *wbcK* gene cloned as a PCR-fragment into pET28	III
pYeO3wbcL	The *wbcL* gene cloned as a PCR-fragment into pET28	III
pYeO3wbcM	The *wbcM* gene cloned as a PCR-fragment into pET28	III
pYeO3wbcQ	The *wbcQ* gene cloned as a PCR-fragment into pET28	III
pTM100wbcO	The *wbcO* gene cloned into pTM100	III
pTM100wbcN	The *wbcN* gene cloned into pTM100	III
pYeO3wbcK-Ala	*wbcK1;*Asp182Ala derivative of pYeO3wbcK	III
pYeO3wbcK-Ser	*wbcK2;*Asp182Ser derivative of pYeO3wbcK	III
pYeO3wbcL-Ala	*wbcL1;*Glu181Ala derivative of pYeO3wbcL	III
pYeO3wbcL-Gln	*wbcL2;*Glu181Gln derivative of pYeO3wbcL	III
pYeO3wbcL-Thr	*wbcL3;*Glu181Thr derivative of pYeO3wbcL	III

Plasmid	Genotype	Reference
pEPlig532	YeO3-532 in pTM100, Tet^R	IV
pEPlig777	YeO3-777 in pTM100, Tet^R	IV
pEPlig1727	YeO3-1727 in pTM100, Tet^R	IV
pEPlig777su	suiside construct of YeO3-777 in pRV1,Clm^R	IV
pEPlig1727su	suiside construct of YeO3-1727 in pRV1, Clm^R	IV
pEPlig532su	suiside construct of YeO3-532 in pRV1, Clm^R	IV
pSW25-lig777del	suiside construct of YeO3-777 in pSW25, Spe^R	IV
pSW29-lig532del	suiside construct of YeO3-532 in pSW29, Kan^R	IV
pMF19	$wbbL_{O16}$ cloned into pEXT21; Spe^R	(Feldman et al, 1999)
pET28a	KanR cloning vector including cleavable N-terminal His-Tag	Novagen
pSW25T	suiside vector, Spe^R	(Demarre et al, 2005)
pSW29	suiside vector, Kan^R	(Demarre et al, 2005)
pRV1	suiside vector, Clm^R	(Skurnik et al, 1995)
pTM100	mobilizable cloning vector, Clm^R Tet^R;	(Michiels & Cornelis, 1991)
pUC18	Cloning vector , Amp^R	(Yanisch-Perron et al, 1985)

a) Amp, ampicillin; Clm, chloramphenicol; Spe, spectinomycin; Tet, tetracycline; Kan, kanamycin.
b) roman numerals refer to original publications
c) see chapter 5.1.4.

4.2. Antibodies, bacteriophages and bacteriocins

Antibodies and bacteriophages used in this work are listed in Table 5. Enterocoliticin, a channel-forming bacteriocin produced by *Y. e.* 29930 (biogroup 1A; serogroup O:7,8 (Strauch et al, 2003)) was kindly provided by Dr. Eckhard Strauch.

Table 5. Antibodies and bacteriophages used in this work.

Antibodies / bacteriophage	Receptor	Used in article(s)
ΦYeO3-12	YeO3 OPS	III
ΦR1-37	YeO3 OC	III,IV
mAb TOM	YeO3 OPS	IV
mAb 2B5	YeO3 OC	III, IV
mAb 898	ECA	IV

4.3. Molecular biology techniques

4.3.1. General DNA techniques (I-IV)

Standard DNA techniques were performed as described (Sambrook & Russel, 2001) and enzymes were used as recommended by the suppliers. To mobilize various plasmid constructs to different *Yersinia* strains triparental conjugation was performed as

described in (Biedzka-Sarek et al, 2005), DNA sequencing reactions were performed in the Sequencing Core of Hartman Institute using Dye Terminator (v3.1) kit (Applied Biosystems). For DNA hybridization the DNA was labeled with DIG High Prime DNA Labeling and Detection Starter kit (Roche).

4.3.2. Construction and mobilization of plasmid pRV16NP

Plasmid pRV16NP was constructed as described (Kiljunen et al, 2005) by introducing the YeO3 OC gene cluster from plasmid pRV7 to pTM100 so that the cloned gene cluster is transcribed from the tetracycline-resistance gene promoter of the vector. pRV16NP was introduced to several *E. coli* strains (C600, NovaBlue(DE3), BL21(DE3)) by electroporation and mobilized to YeO3-c-OC, YeO3-c-OC-R, 8081-c (YeO8) and 8081-c-R2 (YeO8), *Y. p.* KIMD27Nar, *Y. p.* EV-76-c and *Y. pstb.* PBΔwb by triparental conjugation using helper strain HB101/pRK2013. The LPS phenotypes were analyzed by DOC-PAGE analysis.

4.3.3. Construction of the Cat-Mu transposon insertion library (IV)

The Cat-Mu transposon insertion library of YeO3-R1 was constructed by Maria Pajunen at University of Helsinki as described in original publication IV.

4.4. Protein techniques (I, II)

Standard protein chemistry techniques (SDS-PAGE, estimation of the protein amount, purification of histidine (His)-Tagged proteins) were performed with minor modifications as described in original publications.

4.4.1. Enzyme activity assays (I, II)

Preparation of cell extracts and the enzyme activity reactions have been described in original publications (I, II). In publication I the UDP-*N*-acetylglucosamine-4-epimerase activity was determined analyzing the formation of UDP-GalNAc using Dionex high-performance anion-exchange chromatography system combined with pulsed amperometric detection system and the UDP-glucose-4-epimerase activity was measured spectrophotometrically analyzing the formation of UDP-Glc. In publication II the UDP-GlcNAc-4,6-dehydratase activity was determined by analyzing the formation of UDP-2-acetamido-2,6-dideoxy-d-*xylo*-hex-4-ulopyranose (UDP-Sug*p*) using capillary elctrophoresis (CE).

4.5. LPS isolation and analysis (I-IV)

LPS was isolated by proteinase K treatment or small scale hot phenol /water extraction for analysis by gel electrophoresis. Large scale phenol extraction was performed for chemical analysis purposes. Details of the procedures are present in original publications. LPS was analyzed by silver-stained DOC-PAGE as described (Skurnik et al, 1995,

Zhang & Skurnik, 1994), for immunoblotting the DOC-PAGE gels were blotted onto nitrocellulose membranes (Biedzka-Sarek et al, 2005).

Chemical analysis of the LPS structures was performed in Research Center Borstel, Leibniz-Center for Medicine and Biosciences, Borstel, Germany as described in original publications and their supplement data.

4.6. Phage and enterocoliticin techniques

4.6.1. Construction of rough derivatives (III)

To obtain OPS-negative (rough) derivatives of different *Y. e*. O:3 strains I used phage ϕYeO3-12 selection. The desired rough strains were isolated as spontaneous ϕYeO3-12-resistant mutants of corresponding OPS-positive (smooth) strains as described earlier (Skurnik et al, 1995). The OPS-negative phenotypes were confirmed by DOC-PAGE.

4.6.2. ϕR1-37 and enterocoliticin sensitivity assays (II-IV)

Different strains were tested for their ϕR1-37 and enterocoliticin sensitivities by one of the following ways. Either by i) dispersing bacteria onto solid media and applying a drop containing 4×10^3 activity units (AU) of enterocoliticin or 4×10^6 plaque forming units (pfu) of phage or ii) two by pipetting 5 µl of serial dilutions of phage or enterocoliticin stocks (2×10^9 pfu /ml and 2×10^6 AU /ml, respectively) on a freshly grown and dried bacterial lawn. In both ways the formation of a clear lysis zone after 24 h incubation was observed.

The enterocoliticin selection for Cat-Mu transposon library was performed by plating 10^6 colony forming units (CFU) together with 10 AU of enterocoliticin on LA-chloramphenicol (Clm) plates. After 20 h at RT the enterocoliticin resistant colonies were isolated and pure cultures were tested individually to confirm the enterocoliticin resistant phenotype.

4.7. Molecular modeling (I, III)

Molecular modeling was performed by Tiina Salminen in Åbo Akademi as described in original publications I and III in detail.

4.8. Other techniques (III)

Radial diffusion assay was performed by Jose-Antonio Bengoechea at the Laboratory of Microbial Pathogenesis (CSIC), Bunyola, Spain as described in detail in publication III.

5. RESULTS AND DISCUSSION

5.1. Characterization of the biosynthesis of the YeO3 OC (I-IV)

The goal of this PhD study was to elucidate how YeO3 synthesizes the hexasaccharide branch of its LPS called OC. Compared to the classical *E. coli* or *Salmonella* LPS organization, where lipid A -core and OPS are joined so that core interconnects the lipid A to the OPS, the YeO3 LPS is peculiar as both the OC and OPS are linked to IC. When the study was started the YeO3 OC structure was published and its important role in virulence during early phases of infection was known.

As the amount of bacterial isolates resistant to traditional antibiotics is increasing there is constant need for new therapeutics. The detailed understanding of the biosynthesis of virulence factors like YeO3 OC hopefully would provide new knowledge that can help in developing novel antimicrobials. In addition, the different enzymes (particularly, NDP-sugar biosynthesizing enzymes and GTases with characterized linkage specificities) might have wider use both in industry and in applications of basic research.

When I began the study YeO3 OC genes were found to be in a 13-kb cluster localized between *hemH* and *gsk* in the chromosome of this bacterial strain. In other *Yersinia,* the OPS gene cluster is usually found in this location. Also the putative functions for the YeO3 OC gene cluster genes had been postulated according to homology comparisons of their nucleotide sequences with other known genes in the Genbank databases. Until the studies described in this thesis, experimental evidence was missing for all of them.

5.1.1. Characterization of the UDP-N-acetylglucosamine 4-epimerase activity in YeO3 and YeO8 (I)

The first characterized gene in this PhD study was *gne,* formerly named as *galE,* accordingly to YeO3 *gne´*s high sequence similarity to *E. coli galE.* When combining the solved OC structure, sequence analysis of the OC gene cluster and what was known of the outlines of LPS biosynthesis generally it was postulated that two of the nine genes of the OC cluster, *gne* and *wbcP,* are probably involved in the biosynthesis of NDP-sugar precursors UDP-Gal*p*NAc and UDP-Fuc*p*NAc, respectively. According to results from sequence analysis, Gne was predicted to be a soluble protein. However, WbcP possesses transmembrane helices near the N-terminus of the protein and was found to be associated with membrane fragments during the purification procedures. To alleviate the solubility issue of WbcP, a truncated version devoid of the transmembrane helices was constructed, WbcP-S212 (see Section 5.1.2 below for more details). The NDP-sugar precursors for the Gal and Glc residues present in the YeO3 OC are also needed for other purposes in the bacteria and genes responsible for their biosynthesis were thus thought to be most probably located outside of the OC gene cluster.

To test the hypothesis that Gne would be needed for the biosyntheisis of UDP-Gal*p*NAc the possibility to produce UDP-Gal*p*NAc from UDP-Glc*p*NAc in a reaction catalyzed by Gne was tested. A method to identify UDP-Gal*p*NAc and UDP-Glc*p*NAc was also needed. As UDP-Gal*p*NAc and UDP-Glc*p*NAc are epimers they differ only in the conformation of C4 -OH group (Figure 5) and I could not separate these two UDP-sugars from each other by using reverse phase HPLC. Instead I developed a HPLC method with an anion-exchange column using sodium acetate and sodium hydroxide solution as eluents. To test the enzyme activity I set up reactions containing cytoplasmic extracts of either Gne expressing or non-expressing control strain together with UDP-Glc*p*NAc in buffer solution. After incubation at RT, I analyzed the reactions by HPLC.

Using crude cytoplasmic extracts of bacteria I was able to show that both YeO3 *gne* and YeO8 *gne* provide UDP-N-acetylglucosamine 4-epimerase activity. The results obtained with different cell extracts are shown in Figure 2A of original publication I. No UDP-N-acetylglucosamine 4-epimerase activity was present in the host strain *E. coli* C600 extracts; neither UDP-Gal*p*NAc nor UDP-Glc*p*NAc was epimerized. However, the epimerase activity was present in both the C600/pEEgne and C600/pRV28 extracts, telling that both YeO8 Gne and YeO3 Gne have UDP-N-acetylglucosamine 4-epimerase activity. As both UDP-Glc*p*NAc and UDP-Gal*p*NAc functioned as substrates I could show that the reaction is reversible.

Previously, *gne* of YeO3 was named *galE* based on its amino acid sequence similarity to that of *galE* in *E. coli* (Gne and *E. coli* GalE have 53% identity). Thus, *gne* had been proposed to code for UDP-glucose 4-epimerase. In fact I was able to show in a spectrophotometric assay of crude cytoplasmic extracts that YeO8 Gne indeed has in addition to UDP-*N*-acetylglucosamine 4-epimerase activity also UDP-glucose 4-epimerase activity. Furthermore Gne was able to complement *galE* mutant *E. coli* AD9 which supported the previous results.

Biological assay to confirm the UDP-*N*-acetylglucosamine 4-epimerase activity *in vivo* was done by introducing the pEEgne and pRV28 plasmids to *E. coli* C600 /pLZ6010. They both enabled the bacteria to express full length O-ag as expected (see details in article I).

5.1.1.1. Substrate specificity of Gne

To identify the critical Gne residues for substrate specificity the structure of Gne was modeled based on the crystal structures of the human and *E. coli* enzymes. Interestingly human UDP-glucose-4-epimerase (55% identical to Gne) is also able to catalyze both reactions, where as *E. coli* GalE only exhibited UDP-glucose-4-epimerase activity. The unique feature characterizing epimerase reaction is putative rotation of the 4´-ketopyranose within the active site. According to the modeling data there was insufficient space to allow for the rotation of an *N*-acetylated UDP-sugar in *E. coli* GalE, but in YeO8 Gne the situation was opposite. See Figure 3 of article I for details. By site-directed mutagenesis,

I introduced the corresponding bulkier tyrosine (Tyr) residues present in *E. coli* GalE to YeO8 Gne (mutations leusine (Leu)136Tyr, cysteine(Cys)297Tyr and double mutant Leu136Tyr Cys297Tyr) and was able to show that although introducing either of the two Tyr residues was enough to abolish the UDP-N-acetylglucosamine 4-epimerase activity, the UDP-glucose 4-epimerase activity was preserved (Figure 4 of article I).

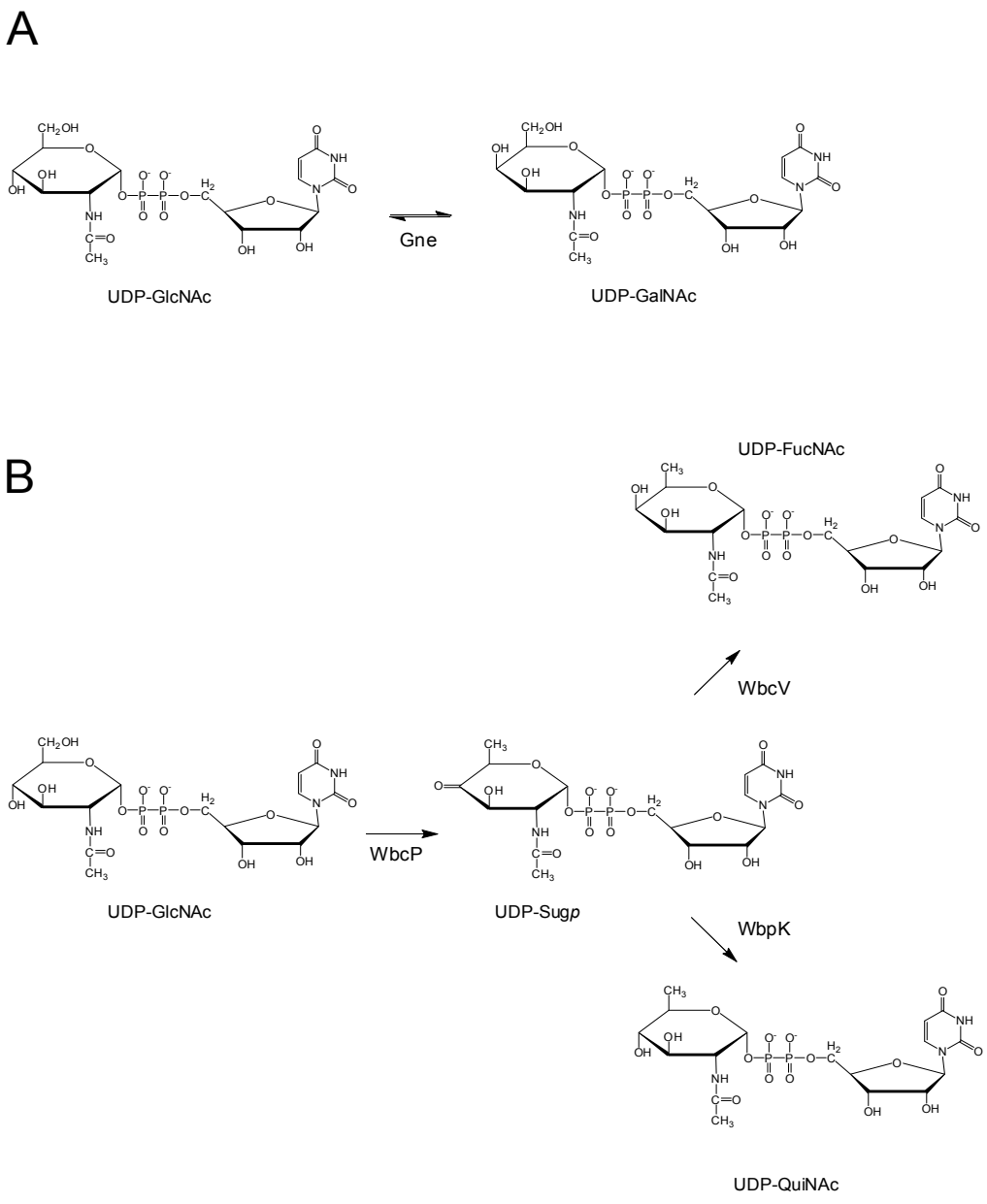

Figure 5. Structures of different UDP-sugars discussed in this work. Enzymes catalyzing the corresponding reactions are also shown.

5.1.2. Functional characterization of the WbcP protein (II)

After characterization of Gne, WbcP was the only candidate of the OC cluster to be responsible for the biosynthesis of the first sugar residue of the OC, previously characterized as FucNAc. The fact that WbcP has 48% identical in amino acid sequence to WbpM, a *P. aeruginosa* enzyme involved in the biosynthesis of UDPFuc*p*NAc, also supported the hypothesis and helped us to design N-terminally truncated versions of WbcP. Similar to the case for WbpM, we hoped that by removing the N-terminal membrane anchor we could purify a soluble but still active version of WbcP.

The basic strategy as for analyzing Gne activity was also applied here. I wanted to demonstrate the WbcP activity by an appearance of a peak in a WbcP catalyzed reaction. The HPLC method used with Gne was not providing sufficient resolution but instead a previously described CE method (see article II for details) was used. Additional challenge was that the expected reaction product (UDP-Fuc*p*NAc) is not available from any commercial source to be purchased and used as a control.

Partially purified cell extract of wt WbcP and purified truncated version of WbcP (WbcP-S212) exhibited enzymatic activity as demonstrated in CE by the consumption of the UDP-Glc*p*NAc substrate peak while the emergence of a new substance became apparent. The new substance was identified as UDP-2-acetamido-2,6-dideoxy-d-*xylo*-hex-4-ulopyranose (UDP-Sug*p*) by the fact that WbgX could use it as its substrate. WbgX is an aminotransferase from *Plesiomonas shigelloides* known to use UDP-Sug*p* to form UDP-2-acetamido-4-amino-2,4,6-trideoxy-D-galactose (Miller, 2006, Shepherd et al, 2000). These results showed that WbcP is a UDP-GlcNAc-4,6-dehydratase that catalyses the reaction from UDP-Glc*p*NAc to UDP-Sug*p*.

5.1.2.1. WbcP is involved in the synthesis of the first OC sugar

In the reaction mixture with partially purified cell extract of WbcP in addition to the UDP-Sug*p*, a second peak emerged in CE. Further analysis identified this compound as UMP and the result supported the hypothesis that WbcP would be involved in the biosynthesis of the first sugar residue of the OC as UMP is released only when the first sugar residue is transferred from UDP-sugar to Und-P. In the following GTase reactions UDP is instead released (for details see the literature review). Appearance of UMP peak also suggested that *E. coli* priming GTase present in the crude membrane preparation could use UDP-Sug*p* as a substrate. This is interesting as there have been discussions of the specificity of the *E. coli* priming transferase WecA concerning UDPGalNAc and UDP-GlcNAc (Rush et al, 2010) as mentioned in the literature review.

5.1.2.2. UDP-Sugp reductase activity in YeO3

In other bacteria, the UDP-Sug*p* is an intermediate in the UDP-Fuc*p*NAc and UDP-Qui*p*NAc synthesis; but to our surprise, in YeO3 the UDP-Sug*p* reductase activity seemed to be missing. This could be due to a tight coupling of the biosynthesis of the UDP-Sug*p*

to the priming transferase or a lack of reductase activity in YeO3. To test the possibilities, I introduced *P. aeruginosa* 4,6-reductases WbpK and WbpV (4,6-reductases in UDP-*N*-acetylquinovosamine (2-acetamido-2,6-dideoxy-d-glucose, Qui*p*NAc) and UDP-Fuc*p*NAc biosynthesis, respectively) to YeO3. The new recombinant strains produce altered LPS. The LPS produced are only partially substituted by OC (Figure 5 of article II) and furthermore instead of synthesizing Sug*p* in the OC, Qui*p*NAc and Fuc*p*NAc, respectively, were found to be incorporated as part of the OC. These results demonstrate that both UDP-Sug*p* reductase activities are missing in YeO3 and when combining with other studies (Forsberg et al, 2003) we can postulate that similar situation occur in all the other organisms that possess Sug*p* in their LPS.

The reason for the partially substituted LPS to be synthesized might be that i) the WbcO (Glycosyltransferase for Sug*p* see article III) is able to add Qui*p*NAc and Fuc*p*NAc to UndP-IC but less efficiently than Sug*p*, ii) WbcQ (GTase for GalNAc see article III) could use UndPP-Qui*p*NAc and/or UndPP-Fuc*p*NAc as substrates, albeit at lower efficiency or iii) either the OC-flippase or OC-ligase is able to transfer or ligate the modified OC less efficiently to lipidA-IC. The first proposal appeared to be consistent with the earlier results with *Rhizobium etli* mutant where UDP-QuiNAc GTase was shown to be able to use UDP-Sug*p* as substrate but less efficiently than UDP-QuiNAc (Forsberg et al, 2003).

5.1.3. Characterization of GTases of the YeO3 OC gene cluster (III)

After solving the biosynthesis of the NDP-sugar precursors for the OC biosynthesis I turned my research focus to ask the following question: how would YeO3 be able to build up the OC hexasaccharide? My hypothesis was that the biosynthesis of OC follows the Wzy-dependent pathway and thus the OC would be synthesized by priming GTase followed by five specific GTases onto an Und-P carrier on the cytoplasmic side of the IM. In the YeO3 OC gene cluster, six genes were postulated to code for the necessary GTases. Using CAZY database to perform analysis of the GTases for YeO3 OC, the putative GTases could be classified to CAZY families (shown in Table 6). The BLAST-search of the five OC GTases against non-redundant protein database sequences revealed significant similarities to many putative GTases; however, there was a general lack of information relating to the functions of the homologues of these GTases and their linkage specificity.

To assign functions for the gene products the strategy was to inactivate these one-by-one by constructing non-polar mutations on every OC gene that is predicted to encode GTases, and then to analyze the phenotypes of the mutated strains. The mutants were expected to produce either no OC or OC with truncated oligosaccharides of different lengths. By using chemical and structural analyses of the mutant LPS the specific functions to primase and five different GTases could be assigned and the exact order in which the transferases build the hexasaccharide could be revealed. The linkage specificities of the primase and the different GTases are listed in Table 6 and Figure 6 summarizes the results.

Table 6. Catalytic specificities of the *Ye* O:3 OC GTases. (III)

Transferase	Donor	Acceptor	Linkage	CaZyGT Family
WbcK	UDP-Glc*p*	Und-P-P-OC5*	β–(1→6)-	GT 2
WbcL	UDP-Glc*p*	Und-P-P-OC4	β–(1→3)-	GT 2
WbcM	UDP-Gal*p*Nac	Und-P-P-OC3	α–(1→6)-	GT 4
WbcN	UDP-Gal*p*	Und-P-P-OC2	α–(1→4)-	GT 4
WbcQ	UDP-Gal*p*Nac	Und-P-P-*Sug*p	α–(1→3)-	GT 4
WbcO	UDP-Sug*p*	Und-P	phosphodiester	--

*OC6 corresponds to wt OC having all six sugar residues, OC5 to OC lacking the β–1-6 linked Glc that is OC having five sugar residues, OC4 to OC having four sugar residues etc.

The GTases characterized in article III provide a valuable supplement to the scanty knowledge concerning linkage specificity of different GTases. To our knowledge there are no GTases with the very same specificity characterized so far. Interestingly the OPS structures of *Salmonella* O66 and *E. coli* O166 were established recently and two linkages found in them are identical to the ones in YeO3 OC. In addition, the functions for corresponding GTase genes were proposed and the results from our experiments was in agreement with their hypothesis that WeiA (WbcL homolog) and WeiC (the WbcN homolog) are Glc-β-(1→3)-GalNAc and Gal-α-(1→4)-GalNAc GTases, respectively. In the future, the hypothesis could be tested with complementation studies by investigating the capability of WeiA to complement *wbcL* of YeO3, and WeiC to complement *wbcN* of YeO3.

5.1.3.1. Catalytic bases of WbcK and WbcL

I find GTases WbcK and WbcL very interesting as they both transfer Glcp from UDP-Glc*p* to Gal*p*NAc but with unique linkage specificities, β–(1→6)-for WbcK and β–(1→3)-for WbcL. In addition, the receiving Gal*p*NAc is different in both cases and the Glc added by WbcL is branching whereas the Glc added by WbcK is terminal. This indicates that the overall three-dimensional space that accommodates the substrates must be different. By sequence comparison and comparative modeling exercises, the catalytic bases of WbcK and WbcL were predicted to be aspartic acid (Asp) 182 and glutamic acid (Glu) 181, respectively. By site-directed mutagenesis I was able to experimentally show that when WbcK-Asp182 was replaced either with alanine (Ala), i.e. D182A, or serine (Ser), i.e. D182S, and when WbcL-Glu181 was replaced with Ala (E181A), glutamine (Gln) (E181Q) or threonine (Thr) (E181T), the corresponding GTase activity was lost. None of the mutated GTases was able to complement the WbcK or WbcL mutant phenotype and thus the catalytic bases are Asp182 for WbcK and Glu181 for WbcL. The results from these experiments indicated the importance of D182 in WbcK activity and E181 in WbcL activity.

5.1.4. Characterization of flippase (unpublished)

If OC biosynthesis follows the Wzy-dependent pathway after the OC is synthesized onto the Und-P carrier it is proposed to be translocated by a flippase to the periplasmic

side of the IM. In the YeO3 OC gene cluster a putative *wzx* gene has been identified and I constructed a site-directed *wzx* YeO3 mutant following the same strategy as for pRV16NP# derived GTase mutants in article III. The *wzx* mutant lacked the OC expression as expected but no further analysis was performed to fully demonstrate the flippase activity of the putative *wzx* gene.

5.1.5. Identification and characterization of the oligosaccharide ligase needed to ligate YeO3 OC to the IC (IV)

After translocation the newly-synthesized OC is ligated to the lipid A IC from the Und-P carrier on the periplasmic side of the IM. As discussed in the literature review there are many enzymes involved but the formation of the IC-OC connecting linkage is proposed to be catalyzed by an enzyme called ligase (ligase is named WaaL in *E. coli*). In the YeO3 OC gene cluster there was no putative ligase gene identified and the whole YeO3 genomic sequence has not been published. The strategy I chose was to isolate a ligase mutant strain. The hypothesis was that an OC ligase mutant would not express OC and that I could therefore isolate ligase mutant candidates using enterocoliticin selection. A random Cat-Mu transposon insertion library of ca. 52000 mutants was constructed to strain YeO3-R1 and in the first round of selection 224 enterocoliticin resistant mutants were isolated.

An enterocoliticin-resistant phenotype could be due to several different reasons. The Cat-Mu transposon insertion could be (i) in the outer core gene cluster, (ii) in gene(s) involved in lipid A or IC biosynthesis, (iii) in an irrelevant gene, therefore the enterocoliticin resistance would be due to spontaneous mutation affecting the OC biosynthesis, (iv) in the ligase gene, or (v) in any other gene that directly or indirectly could cause enterocoliticin resistance.

When characterizing mutants from the Cat-Mu transposon library representatives from most of the above mentioned categories (i, ii, iii and v) were identified. In deep rough mutant strains the Cat-Mu transposon insertions were identified in genes that are apparently involved in IC biosynthesis, some of which code for putative GTases (*waaF, waaC, waaE)* or are needed for biosynthesis of NDP-sugar-precursors (*galU)*. One regulation mechanism of YeO3 OC biosynthesis became obvious via the Cat-Mu transposon insertion mutants. Many OC lacking mutants had Cat-Mu transposon in the *rfaH* gene. RfaH is transcription elongation factor and potentially together with other factors modifies the RNA polymerase complex to increase its processivity and allowing transcription to proceed over long distances (Belogurov et al, 2010). As OC genes are clustered in the 13 kb cluster, we propose that RfaH allows efficient expression of the OC cluster genes.

In strain YeO3-R1-M089 Cat-Mu transposon was inserted in a gene that likely would not be directly involved in OC biosynthesis. The gene homolog in YeO8 strain 8081 is *ye1725*. Examining the gene organization around the *ye1725* gene we noticed that it is a first gene in a small operon that raised a possibility that enterocoliticin resistance could be due to polar effect on any of the downstream genes. Homology searches and prediction

of membrane topology indicated that the product of gene *ye1727* had characteristics that could be associated with a ligase. A search of the partial genomic sequence data of YeO3 showed that a full homolog of *ye1727* was present in YeO3 genome. We decided to study in more detail this gene that we later named as $waaL_{os}$.

A putative *waaL* gene in *Y. p.* (*ypo0417*) had been identified (Anisimov et al, 2010, Knirel et al, 2006). The *ypo0417* homologue in the YeO8 genome is *ye0532*; and a homologue was also identified from the partial genomic sequence of YeO3 encoding a protein of 414 residues. This gene, later named as $waaL_{ps}$, was also selected for more detailed studies. Finally, a third ligase candidate gene (*ye0777*) from YeO8 genome was identified in a BLAST search with the YeO3 $WaaL_{os}$ and $WaaL_{ps}$ amino acid sequences. A homolog to this gene, later named as $waaL_{eca}$, was also found from the partial genomic sequence of YeO3 encoding a 554 amino acid protein.

When $waaL_{os}$, $waaL_{eca}$ and $waaL_{ps}$ genes of YeO3 were individually introduced to Δ*waaL E. coli* all the constructed strains expressed OPS. Thus the genes complemented the *waaL* mutant phenotype and I could conclude that $waaL_{os}$, $waaL_{eca}$ and $waaL_{ps}$ of YeO3 code for ligase activity in *E. coli* background.

Silver stained DOC-PAGE and immunoblotting analysis of the LPS phenotypes of single, double and triple mutants of $waaL_{os}$, $waaL_{ps}$ and $waaL_{eca}$ genes in YeO3 revealed that $waaL_{op}$ codes for OC ligase and $waaL_{ps}$ for OPS ligase (Figures in article IV). We have obtained results from immunoblot analysis of the ECA expression indicating that $waaL_{eca}$ codes for ECA ligase but further analysis is still required.

While the single and double mutants expressed significantly reduced and varying amounts of OC and OPS, only the triple mutant completely lacked OC, OPS and ECA indicating that the ligases are somewhat promiscuous with regards to the transferred oligo- or polysaccharide unit.

Until now WaaL has been the only enzyme associated to the ligation of O-antigen to the lipid A-core. It has been proposed that the specific acceptor structure (lipid A core) and the Und-P carrier of the donor rather than the sugar part would be required for ligase activity (Guo et al, 2008, Heinrichs et al, 1998b). The special situation of YeO3 where both OC and OPS are ligated to lipidA IC provides an interesting case and its need for different ligases becomes obvious. Our results show that indeed there are different ligases available but without further biochemical studies we cannot explain the detailed reaction chemistry or the substrate requirements.

It would be really interesting to solve the 3D structures of the different YeO3 ligases and compare them to the structure of the *E. coli* WaaL that they all were able to complement. Maybe this comparison could help us understand the critical differences behind the different specificities of $WaaL_{os}$, $WaaL_{ps}$ and $WaaL_{eca}$ and would also explain what would be the critical elements in their structural similarities that allow them to partially complement each other.

5.1.6. Sequence based postulations need to be experimentally proven

The major aim of the work was to obtain experimental evidence to specifically assign all the enzymes needed for the biosynthesis of the OC. It seems that most of the sequence based predictions were correct especially when concerning the predictions of enzymes belonging to different families of biosynthetic enzymes but it became evident that according to sequence analysis the detailed specificities can only be postulated and experimental evidence is needed to support them.

5.1.6.1. Postulating specificities for NDP-sugar biosynthesizing enzymes

In the case of NDP-sugar biosynthesizing enzymes Gne was previously predicted to catalyze the reaction from UDP-GlcNAc to UDP GalNAc. I was able to show that indeed Gne has UDP-*N*-acetylglucosamine-4-epimerase activity but that it also has UDP-glucose-4-epimerase activity. By using an approach of modeling together with site-directed mutagenesis of Gne and the comparison of sequences of proteins having over 30% identity to Gne had allowed us to postulate that enzymes that have at their respective residue positions 136 and 297 either phenylalanine (Phe)-Thr, Phe-Tyr, Tyr-Phe, Tyr-serine (Ser) or Tyr-Tyr will most probably have only the UDP-glucose-4-epimerase activity whereas the other combinations (with smaller amino acid residues and thus space for rotation see article I for details) will probably have only the UDP-*N*-acetylglucosamine-4-epimerase activity or both. These predictions warrant future studies to confirm them and a simple biological assay to get the first idea of the activity of a Gne homologue with unknown function is proposed in article I.

The precursor for the first sugar unit in the OC (previously characterized as Fuc*p*NAc) was initially predicted to be biosynthesized from UDP-Gal*p*NAc in a reaction catalyzed by WbcP. The results of article II demonstrated that in fact UDP-Gal*p*NAc is not the substrate for WbcP but instead UDP-Glc*p*NAc is. It also became evident that the reaction product is UDP-Sug*p* and the result was in good agreement with the revised structure of YeO3 OC provided in article II. Also these results demonstrated that nucleotide sequence-based predictions are insufficient and experimental confirmations are necessary. The Sug*p* is not commonly found in LPS structures and it had never been identified as a constituent of *Yersinia* cell envelope before. Interestingly the recently sequenced LPS gene cluster of *Yersinia* sp. strain Å125 KOH12 which encodes highly identical homologues of WbcO (the priming transferase of YeO3, see article III) and WbcP and it is plausible that the Å125 LPS also carries Sug*p*. By introducing the Å125 WbcP to WbcP mutant YeO3 the activity could be tested and defined.

5.1.6.2. Predicting GTase specificities

The second type of enzyme characterized in this work was the GTases. The enormous sequence data available allows predictions of subfamilies and functions for different enzymes but especially in the case of GTases experimental evidence is needed to conclusively specify their functions. Our results in article III suggest that the initial

prediction for the priming transferase to be WbcO is most probably correct but direct experimental evidence is still missing as for an *in vitro* experiment we did not have the reagents, UDP-Sug*p* as donor and Und-P (or equivalent) as acceptor, available.

The experimentally proven GTase specificities of article III of the remaining five GTases agree with the initial predictions but only to some extent. Thus the importance of the experimental evidence is emphasized and a biological complementation assay used to verify the mutations in article III to test other GTases predicted to have the same specificities is suggested.

5.1.6.3. Sequence-based predictions used to identify the ligase gene(s)

When characterizing the YeO3 OC ligase the sequence-based predictions to find the ligase gene candidates were used. In these *in silico* analyses, the different prediction programs available in the web-based servers have been very useful. Interestingly in YeO3 there are at least three different putative ligase enzymes involved in the ligations of different oligo- or polysaccharides to membrane bound carbohydrate acceptors. All these enzymes exhibited ligase activity when used to complement an *E. coli waaL* mutant but they seemed to have different specificities in YeO3 even though they were able to complement each other to some extent. This could explain why the use of the more traditional way to find the ligase using Cat-Mu transposon insertion library was not successful due to the overlap functions.

5.1.7. Conclusions on the characterization of the OC biosynthesis of YeO3

As a conclusion, this study defined the biosynthesis of YeO3 OC. All the results supported the hypothesis that YeO3 OC synthesis follows the Wzy-dependent pathway even though the pathway used was not impugned. By providing experimental evidence I was able to define the activities of nine different gene products and together they assign a bundle of information of the coordinated synthesis of YeO3 OC. This new knowledge can be helpful in assigning functions for other gene products involved in the syntheses of different carbohydrates. In addition it can be helpful in understanding the overall picture of the very diverse biosynthesis scheme of the different bacterial virulence factors.

5.2. Characterization of OC as a surface receptor (II, III, unpublished)

The second part of this study illustrated the characteristics of YeO3 OC as a surface structure. YeO3 OC has been proposed to act as a receptor for bacteriophage ϕR1-37 (Skurnik et al, 1995) and enterocoliticin (Strauch et al, 2003). YeO3 OC is also recognized by a monoclonal antibody (mAb) 2B5 (Biedzka-Sarek et al, 2005, Pekkola-Heino et al, 1987).

To further characterize the phage receptor I showed that the phage recognition of bacteria was abrogated when the bacteria were treated with periodate, suggesting that the receptor is a carbohydrate. Boiling and proteinase K treatment did not destroy the

receptor, indicating that receptor contains no protein structures. To confirm that the OC is the receptor, I cloned the YeO3 OC gene cluster into a mobilizable vector-construct. The plasmid was named pRV16NP and when this was transformed into an OC deletion mutant strain YeO3-c-OC, it restored the production of OC.

When I mobilized the plasmid pRV16NP into different ϕR1-37 resistant and OC negative *Yersinia* and *E. coli* strains they all expressed the YeO3 OC (Figure 4 in (Kiljunen et al, 2005)) and became ϕR1-37 sensitive. These results suggested that YeO3 OC is the receptor of ϕR1-37.

To further characterize the minimal requirement for the ϕR1-37 recognition and to see if it differs from the requirements for enterocoliticin and mAb 2B5 recognition I tested the strains expressing truncated OC of different lengths. Indeed the requirements were different and they have been summarized in Figure 6.

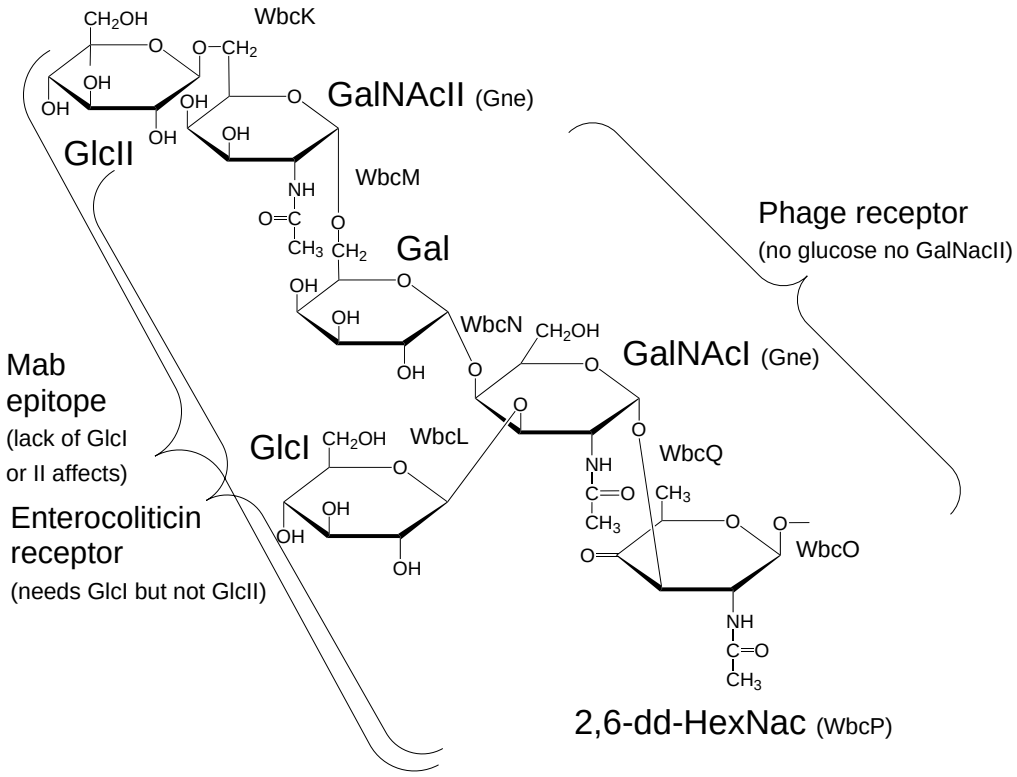

Figure 6. Biosynthesis of YeO3 OC and the different requirements for the ϕR1-37, enterocoliticin and mAb 2B5 recognitions.

5.2.1. Is the OC all that is required?

I have not yet been able to definitely show that the YeO3 OC alone is enough for ϕR1-37 recognition. It might be that some deeper LPS parts (mainly from the IC) might also be

involved. The problem to test this stems from the fact that I have not been able to express OC in deep rough mutants. Apparently for the ligase to add the OC to the lipid-A IC there must be correct kind of acceptor available. And that is missing from the deep rough mutants. The (inner) core structures of the strains were YeO3 OC has been expressed (*E. coli*, YeO8, YeO3, *Y. p.* and *Y. pstb.*) are not identical but maybe there is still enough similarities to provide a part for the ϕR1-37 receptor.

Interestingly a clear difference in phage sensitivity behavior compared to enterocoliticin sensitivity was seen; the ϕR1-37 sensitive *E. coli* NovaBlue/pRV16NP was resistant to enterocoliticin. This could be due to at least two reasons i) enterocoliticin uses YeO3 IC as part of its receptor and that is missing from *E. coli* expressing OC or ii) as the *E. coli* LPS is not fully substituted by the OC, the density of OC is not high enough for the enterocoliticin. The first reason is more likely to be true as the LPS of *wbcK* YeO3 is neither fully substituted by the truncated OC but the strain is still enterocoliticin sensitive.

5.2.2. Is Sug*p* a part of the receptor?

To know if the rare keto-sugar (Sug*p*) present in the OC would be important for the ϕR1-37, enterocoliticin and/ or mAb2B5 recognition the neutralization potential of natural and hydrazine-treated LPS was tested. Mild hydrazine treatment of YeO3-R1 LPS quantitatively reduces the Sug*p* to Qui*p*NAc (and the only additional difference in the hydrazine treated LPS is the loss of the two O-linked fatty acids, see article II for details). When compared to wt LPS, 100 and 1000 times more hydrazine-treated LPS was not able to neutralize either ϕR1-37 or enterocoliticin, respectively. In a dot-blot assay 400 times more hydrazine-treated LPS did not interact with the mAb 2B5 when compared to wt LPS. These results showed that Sug*p* is important for such interactions, i.e. it either forms part of the receptor structure of both ϕR1-37 and enterocoliticin, and the epitope for recognized by mAb 2B5 or it allows the OC to take the correct 3D structure that QuiNAc would prevent. It is worth mentioning that in this experiment the OC substitution of the LPS was not a problem as wt LPS was used.

To test if the presence of Qui*p*NAc or Fuc*p*NAc in the OC structure has any consequences *in situ* on bacterial surface I compared YeO3-R1 to the recombinant strains expressing WbpV or WbpK (YeO3-R1/pEPwbpV and YeO3-R1/pEPwbpK, respectively). The recombinant strains showed altered but different sensitivities. Strain expressing Qui*p*NAc (YeO3-R1/pEPwbpV) was less sensitive to ϕR1-37 infection and strains expressing Fuc*p*NAc (YeO3-R1/pEPwbpK) were fully sensitive. Both recombinant strains were highly resistant to enterocoliticin. These sensitivity changes could be a direct effect of the substitution of Sug*p* by Qui*p*NAc or Fuc*p*NAc in the OC structure or due to decreased expression of the modified OC (Figure 5 of article II). In the case of ϕR1-37 the LPS density is most probably not the reason (see results with OC expressing *E. coli)* and I would suggest that substituting Sug*p* to QuiNAc has influence on the phage recognition (the small sensitivity might be due to the minor Sug*p* containing portion,

see mAb analysis) but that most probably modified OC where Sug*p* has been substituted with FucNAc could act as a phage receptor. The hydrazine treatment results do not argue against this hypothesis since there the Sug*p* has been modified to QuiNAc.

In the case of sensitivity to enterocoliticin, the density of LPS seems not either to be important (see discussion in 5.2.1.) and thus the resistance to could be a direct effect of the substitution of Sug*p* by Qui*p*NAc or Fuc*p*NAc.

Monoclonal antibody 2B5 recognized the OC-containing band of the recombinant strains in immunoblotting which suggested that either the Sug*p* is not crucial for the mAb 2B5 epitope or the *Sugp* is present as a minor constituent in the OC band. The latter is more likely since after chemical analysis (see article II for details), it was estimated that up to 5-10% of the LPS synthesized by these bacteria indeed contained the Sug*p,* an amount sufficient for the highly sensitive immunoblotting method.

In conclusion, Sug*p* is crucial for the biological role of OC, that is, in the ϕR1-37 and enterocoliticin receptor structure and in the epitope of mAb 2B5. This sugar is most probably crucial also for the virulence of YeO3; however, at present, this hypothesis cannot be tested until a strain expressing LPS that would be fully substituted by Qui*p*NAc or Fuc*p*NAc containing OC has been engineered.

5.2.3. Conclusions from the receptor recognition

Bacteriophages, bacteriocins and mAbs are widely used as microbiological tools. Using phages and bacteriocins, different mutants can be isolated; and mAbs are often used to differentiate some minor structural differences that are difficult to distinguish with standard techniques such as SDS-PAGE for resolving LPS of different sizes. The structural requirements for the YeO3 OC recognition by ϕR1-37, enterocoliticin and mAb2B5 were shown to be different and their detailed characterization is helpful when using these microbiological tools in the future. They also demonstrate the strict demands for the specific recognitions. Often when phage receptor structures are investigated mAbs against the structures are used for analytical purposes. As our results demonstrate, the fact that the recognition specificities of phages and mAbs specific for the same receptor do not exactly overlap may cause problems when widening the use of the phage and corresponding mAb to analyze the cell surfaces of different strains, *e.g.* for typing purposes.

6. SUMMARY

In this work I was able to delineate the way how YeO3 is able to synthesize OC. This includes the characterization of three different kinds of enzymes involved in the biosynthesis of the YeO3 OC.

I showed that Gne and WbcP are enzymes that are required for the biosynthesis of the correct UDP-sugar precursors. Gne is a UDP-*N*-acetylglucosamine 4-epimerase and thus provides UDP-GalNAc for the biosynthesis. WbcP is a UDP-GlcNAc-4,6-dehydratase and its enzymatic reaction yields UDP-2-acetamido-2,6-dideoxy-d-*xylo*-hex-4-ulopyranose as the product.

The second type of enzymes characterized in this work is GTases. I was able to characterize linkage specificities of five different UDP-sugar glycosyltrasferases and the priming transferase needed for the OC hexasaccharide to be synthesized on Und-P lipid carrier. The bottom-up sequential order of building the OC is due to the activities of WbcO, WbcQ, WbcN, WbcM, WbcL and WbcK. The linkage specificities conferred by each of these GTases have been listed in Table 6.

Characterization of the OC ligase represents the third type of enzyme. Interestingly we found out that in YeO3 there are different ligases for OC, O-ag and ECA. $WaaL_{os}$ adds OC onto the lipid A-IC and $WaaL_{ps}$ adds O-ag to the lipid A-IC. $WaaL_{eca}$ is most probably involved in the ligation of ECA.

In this work I also showed that YeO3 OC acts as a receptor for ϕR1-37 and that the minimum requirements for the recognition of the YeO3 OC by the ϕR1-37, enterocoliticin and mAb 2B5 are different. My results demonstrated that Sug*p* is crucial for the biological role of OC.

All together this work provides an extensive overview of the biosynthesis of YeO3 OC as it provides a substantial amount of information of the stepwise and coordinated synthesis of the *Ye* O:3 OC hexasaccharide and detailed information of its properties as a receptor.

ACKNOWLEDGEMENTS

This study was carried out at the Department of Medical Biochemistry and Genetics, University of Turku; the Infection Biology research Program, the Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki, and the Glycoscience Graduate School, Finland. I am most grateful to each person who gave me support, advice or encouragement during the work and made the present achievement possible. In particular, I want to thank the following individuals:

Professors Jukka Finne and Klaus Elenius, the former and present head, respectively, of the Department of Medical Biochemistry and Genetics, University of Turku, and professor Seppo Meri, the head of the Infection Biology Research Program, the Department of Bacteriology and Immunology, Haartman Institute, University of Helsinki, who have provided me with the research facilities for this study and been most supportive.

I wish to express my deepest gratitude to my supervisors, Professors Mikael Skurnik and Jose-Antonio Bengoechea. Miikki`s whose sincere enthusiasm and positive attitude towards science are transmitted to people around him. I feel that he has clearly inspired me and set great examples to the way I do my research. With Josean we spent memorable times together and he taught me many things while he was working as a post-doc in Miikki´s laboratory. As a co-supervisor his comments and help concerning my projects and the preparation of publications are greatly acknowledged.

The Glycoscience Graduate School is acknowledged for offering interesting courses and for providing funding for trips and more importantly, for providing a forum for meeting so many wonderful new people.

Professors Miguel Valvano and Elke Schweda are kindly acknowledged for reviewing the thesis and valuable comments. I thank my supervisory committee Docent Eric Bertoft, Professor Jukka Finne and Professor Risto Renkonen for their interest in my work. Eric is also acknowledged for all the help and pleasant discussions while I was using Dionex in his laboratory.

I would also like to thank Professor Joseph S. Lam both for reviewing the English language of the thesis, for the fruitful collaboration, and friendship since my visit to his laboratory. Professor Yuriy A. Knirel, Professor Maria Fredriksson-Ahomaa and Dr. Katarzyna Anna Duda are kindly acknowledged for sharing their expertice while I was finishing writing of this thesis.

MSc. Julia Batzilla, Dr. Sabine Borowski, Dr. Katarzyna Anna Duda, Dr. Anna Hanuszkiewicz, Professor Otto Holst, Dr. Heidi Hyytiäinen, Dr. Zbigniew Kaczynski, Dr. Katarzyna Kasperkiewicz, MSc. Zhilin Li, Dr. Buko Lindner, Dr. Wayne L. Miller, Docent Maria Pajunen, Docent Zofia Piotrowska-Seget, Professor Joanna Radziejewska-Lebrecht, Dr. Alexander Rakin, Dr. Tiina Salminen, Mrs. Reija Venho and Professor Christian Vogel, my co-authors in the original publications of this thesis are all cordially acknowledged

for contributing their results to make the papers significant contributions to the field of glycobiology. Because of the multi-disciplinary nature of the research, I couldn't have been able to finish these papers at the calibre of science that was shown in them.

All the past and present members of the Yersinia Research Group, especially Dr. Marta Biezka-Sarek, Dr. Yasmin El Tahir, Mr. Jukka Karhu, Dr. Saija Kiljunen, Docent Maria Pajunen, Ms. Anne Peippo, and Mrs. Reija Venho have had great influence on the success of this work. Working with you has been a joy and you have taught me many practical laboratory skills that I would never have been able to learn from text books.

I also want to acknowledge all the former and present people working at the Department of Medical Biochemistry and Genetics for creating a friendly atmosphere, especially Mrs. Rakel Mattsson, Mrs. Liisa Peltonen, and Mrs. Raili Salonen have helped me to deal with many practical problems. After the Yersinia Group had gradually moved to the University of Helsinki over a period of time, I was privileged to be "adopted" by Finne´s group, and especially, I'd like to thank Mrs. Elina Jakobsson (MSc), Dr. Anne Jokilammi, Mrs. Annika Kouki (MSc), Docent Vuokko Loimaranta and Docent Sauli Haataja for all the help and fun times concerning science and other aspects of life. Anne´s help in all practical matters concerning the preparation to the defence and Vuokko´s critical comments on my manuscripts have been invaluable.

They say that it needs a whole village to raise a child. I feel that to me "the village" who provided strengths and support are actually my friends and relatives. I want to thank all the friends and acquaintances from guide troop Hakkistytöt as well as from SELKU ry. The experiences in putting things together and feeling the joy of success after difficulties and hard work, as well as gaining the experience that everyone is valuable as himself/herself provides good basics to know what life is all about.

In addition, I was lucky and got involved in a great team already in day care. I want to thank LLM Pauliina Kanerva, Mrs. Hanna Kivisaari-Comas (MA), and Dr. Laura Soupas for your friendship and all the moments we have had together over the past thirty years.

Finally, this thesis is dedicated to my family; my late mother Matleena, whose values have guided me through life, my father Risto, little sister Sini-Päivikki, and Grandmother Inga, whose love, support, and belief in me that I can feel every day, as well as to my dear husband Sami and our three perfect treasures: Isla, Hugo and Vili-Petteri. You are the symbols of happiness and remind me what is really important in life. I feel really lucky and privileged to be your wife and mother, respectively.

This study was financially supported by Glycoscience Graduate School (GGS), Academy of Finland, the Juselius Foundation, Biocity Turku Centre of Excellence, the Magnus Ehrnrooth Foundation and the Medical Faculty of the University of Turku.

Turku, October 2010

Elina Pirttimaa

REFERENCES

Abeyrathne PD, Daniels C, Poon KK, Matewish MJ, & Lam JS (2005) Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of Pseudomonas aeruginosa lipopolysaccharide. *J Bacteriol* **187**: 3002-3012

Abeyrathne PD & Lam JS (2007) WaaL of Pseudomonas aeruginosa utilizes ATP in in vitro ligation of O antigen onto lipid A-core. *Mol Microbiol* **65**: 1345-1359

Achtman M, Morelli G, Zhu P, Wirth T, Diehl I, Kusecek B, Vogler AJ, Wagner DM, Allender CJ, Easterday WR, Chenal-Francisque V, Worsham P, Thomson NR, Parkhill J, Lindler LE, Carniel E, & Keim P (2004) Microevolution and history of the plague bacillus, Yersinia pestis. *Proc Natl Acad Sci U S A* **101**: 17837-17842

Achtman M, Zurth K, Morelli G, Torrea G, Guiyoule A, & Carniel E (1999) Yersinia pestis, the cause of plague, is a recently emerged clone of Yersinia pseudotuberculosis. *Proc Natl Acad Sci U S A* **96**: 14043-14048

Acker G, Knapp W, Wartenberg K, & Mayer H (1981) Localization of enterobacterial common antigen in Yersinia enterocolitica by the immunoferritin technique. *J Bacteriol* **147**: 602-611

al-Hendy A, Toivanen P, & Skurnik M (1992) Lipopolysaccharide O side chain of Yersinia enterocolitica O:3 is an essential virulence factor in an orally infected murine model. *Infect Immun* **60**: 870-875

al-Hendy A, Toivanen P, & Skurnik M (1991a) The effect of growth temperature on the biosynthesis of Yersinia enterocolitica O:3 lipopolysaccharide: temperature regulates the transcription of the rfb but not of the rfa region. *Microb Pathog* **10**: 81-86

al-Hendy A, Toivanen P, & Skurnik M (1991b) Expression cloning of Yersinia enterocolitica O:3 rfb gene cluster in Escherichia coli K12. *Microb Pathog* **10**: 47-59

Anisimov AP, Dentovskaya SV, Kondakova AN, Lindner B, Shaikhutdinova RZ, Kocharova NA, Senchenkova SN, and Knirel YA (2010) Yersinia pestis Lipopolysaccharide in Host-Pathogen Interactions. In *The Challenge of Highly Pathogenic Microorganisms. Mechanisms of Virulence and Novel Medical Countermeasures,* Shafferman A, Ordentlich A, and Velan B (eds) pp 77-87. Springer: New York

Appleyard RK (1954) Segregation of New Lysogenic Types during Growth of a Doubly Lysogenic Strain Derived from Escherichia Coli K12. *Genetics* **39**: 440-452

Aussel L, Therisod H, Karibian D, Perry MB, Bruneteau M, & Caroff M (2000) Novel variation of lipid A structures in strains of different Yersinia species. *FEBS Lett* **465**: 87-92

Babic A, Guerout AM, & Mazel D (2008) Construction of an improved RP4 (RK2)-based conjugative system. *Res Microbiol* **159**: 545-549

Bachmann BJ (1972) Pedigrees of some mutant strains of Escherichia coli K-12. *Bacteriol Rev* **36**: 525-557

Barr K, Klena J, & Rick PD (1999) The modality of enterobacterial common antigen polysaccharide chain lengths is regulated by o349 of the wec gene cluster of Escherichia coli K-12. *J Bacteriol* **181**: 6564-6568

Bastin DA, Stevenson G, Brown PK, Haase A, & Reeves PR (1993) Repeat unit polysaccharides of bacteria: a model for polymerization resembling that of ribosomes and fatty acid synthetase, with a novel mechanism for determining chain length. *Mol Microbiol* **7**: 725-734

Belogurov GA, Sevostyanova A, Svetlov V, & Artsimovitch I (2010) Functional regions of the N-terminal domain of the antiterminator RfaH. *Mol Microbiol* **76**: 286-301

Bengoechea JA, Najdenski H, & Skurnik M (2004) Lipopolysaccharide O antigen status of Yersinia enterocolitica O:8 is essential for virulence and absence of O antigen affects the expression of other Yersinia virulence factors. *Mol Microbiol* **52**: 451-469

Bengoechea JA & Skurnik M (2000) Temperature-regulated efflux pump/potassium antiporter system mediates resistance to cationic antimicrobial peptides in Yersinia. *Mol Microbiol* **37**: 67-80

Bengoechea JA, Zhang L, Toivanen P, & Skurnik M (2002) Regulatory network of lipopolysaccharide O-antigen biosynthesis in Yersinia enterocolitica includes cell envelope-dependent signals. *Mol Microbiol* **44**: 1045-1062

Biedzka-Sarek M, Venho R, & Skurnik M (2005) Role of YadA, Ail, and Lipopolysaccharide in Serum Resistance of Yersinia enterocolitica Serotype O:3. *Infect Immun* **73**: 2232-2244

Bogdanovich T, Carniel E, Fukushima H, & Skurnik M (2003) Use of O-antigen gene cluster-specific PCRs for the identification and O-genotyping of Yersinia pseudotuberculosis and Yersinia pestis. *J Clin Microbiol* **41**: 5103-5112

Boghenbor KK, On SL, Kokotovic B, Baumgartner A, Wassenaar TM, Wittwer M, Bissig-Choisat B, & Frey J (2006) Genotyping of human and porcine Yersinia enterocolitica, Yersinia intermedia, and Yersinia bercovieri strains from Switzerland by

amplified fragment length polymorphism analysis. *Appl Environ Microbiol* **72**: 4061-4066

Bos MP, Tefsen B, Geurtsen J, & Tommassen J (2004) Identification of an outer membrane protein required for the transport of lipopolysaccharide to the bacterial cell surface. *Proc Natl Acad Sci U S A* **101**: 9417-9422

Bottone EJ (1999) Yersinia enterocolitica: overview and epidemiologic correlates. *Microbes Infect* **1**: 323-333

Bottone EJ (1997) Yersinia enterocolitica: the charisma continues. *Clin Microbiol Rev* **10**: 257-276

Bouhss A, Trunkfield AE, Bugg TD, & Mengin-Lecreulx D (2008) The biosynthesis of peptidoglycan lipid-linked intermediates. *FEMS Microbiol Rev* **32**: 208-233

Breton C, Snajdrova L, Jeanneau C, Koca J, & Imberty A (2006) Structures and mechanisms of glycosyltransferases. *Glycobiology* **16**: 29R-37R

Brubaker RR (1991) Factors promoting acute and chronic diseases caused by Yersiniae. *Clin Microbiol Rev* **4**: 309-324

Burk O & Klempnauer KH (1991) Estrogen-dependent alterations in differentiation state of myeloid cells caused by a v-myb/estrogen receptor fusion protein. *EMBO J* **10**: 3713-3719

Burrows LL, Chow D, & Lam JS (1997) Pseudomonas aeruginosa B-band O-antigen chain length is modulated by Wzz (Ro1). *J Bacteriol* **179**: 1482-1489

Butler T (1983) *Plague and Other Yersinia Infections.* Plenum Publishing Corp: 233 Spring Street, New York, NY 10013.

Campbell JA, Davies GJ, Bulone V, & Henrissat B (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J* **326 (Pt 3)**: 929-939

Carniel E, Autenrieth I, Cornelis G, Fukushima H, Guinet F, Isberg R, Pham J, Prentice M, Simonet M, Skurnik M, Wauters G (2006) Y. enterocolitica and Y. pseudotuberculosis. In *Prokaryotes,* Anonymous pp 270-398.

Carter JA, Jimenez JC, Zaldivar M, Alvarez SA, Marolda CL, Valvano MA, & Contreras I (2009) The cellular level of O-antigen polymerase Wzy determines chain length regulation by WzzB and WzzpHS-2 in Shigella flexneri 2a. *Microbiology* **155**: 3260-3269

Carty SM, Sreekumar KR, & Raetz CR (1999) Effect of cold shock on lipid A biosynthesis in Escherichia coli. Induction At 12 degrees C of an acyltransferase specific for palmitoleoyl-acyl carrier protein. *J Biol Chem* **274**: 9677-9685

Chen PE, Cook C, Stewart AC, Nagarajan N, Sommer DD, Pop M, Thomason B, Kiley MP, Lentz S, Nolan N, Sozhamannan S, Sulakvelidze A, Mateczun A, Du L, Zwick ME, & Read TD (2010) Genomic characterization of the Yersinia genus. *Genome Biol* **11**: R1

Clarke BR, Cuthbertson L, & Whitfield C (2004) Nonreducing terminal modifications determine the chain length of polymannose O antigens of Escherichia coli and couple chain termination to polymer export via an ATP-binding cassette transporter. *J Biol Chem* **279**: 35709-35718

Coutinho PM, Deleury E, Davies GJ, & Henrissat B (2003) An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* **328**: 307-317

Creuzenet C, Belanger M, Wakarchuk WW, & Lam JS (2000) Expression, purification, and biochemical characterization of WbpP, a new UDP-GlcNAc C4 epimerase from Pseudomonas aeruginosa serotype O6. *J Biol Chem* **275**: 19060-19067

Cunneen MM, De Castro C, Kenyon J, Parrilli M, Reeves PR, Molinaro A, Holst O, & Skurnik M (2009) The O-specific polysaccharide structure and biosynthetic gene cluster of Yersinia pseudotuberculosis serotype O:11. *Carbohydr Res* **344**: 1533-1540

D Souza JM, Samuel GN, & Reeves PR (2005) Evolutionary origins and sequence of the Escherichia coli O4 O-antigen gene cluster. *FEMS Microbiol Lett* **244**: 27-32

Daniels C, Griffiths C, Cowles B, & Lam JS (2002) Pseudomonas aeruginosa O-antigen chain length is determined before ligation to lipid A core. *Environ Microbiol* **4**: 883-897

Daniels C & Morona R (1999) Analysis of Shigella flexneri wzz (Rol) function by mutagenesis and cross-linking: wzz is able to oligomerize. *Mol Microbiol* **34**: 181-194

Daniels C, Vindurampulle C, & Morona R (1998) Overexpression and topology of the Shigella flexneri O-antigen polymerase (Rfc/Wzy). *Mol Microbiol* **28**: 1211-1222

Deadman ME, Lundstrom SL, Schweda EK, Moxon ER, & Hood DW (2006) Specific amino acids of the glycosyltransferase LpsA direct the addition of glucose or galactose to the terminal inner core heptose of Haemophilus influenzae lipopolysaccharide via alternative linkages. *J Biol Chem* **281**: 29455-29467

DebRoy C, Roberts E, Kundrat J, Davis MA, Briggs CE, & Fratamico PM (2004) Detection of Escherichia coli serogroups O26 and O113 by PCR amplification of the wzx and wzy genes. *Appl Environ Microbiol* **70**: 1830-1832

Demarre G, Guerout AM, Matsumoto-Mashimo C, Rowe-Magnus DA, Marliere P, & Mazel D (2005) A new family of mobilizable suicide plasmids based on broad host range R388 plasmid (IncW) and RP4 plasmid (IncPalpha) conjugative machineries

and their cognate Escherichia coli host strains. *Res Microbiol* **156**: 245-255

Devignat R (1951) Varieties of Pasteurella pestis; new hypothesis. *Bull World Health Organ* **4**: 247-263

Ditta G, Stanfield S, Corbin D, & Helinski DR (1980) Broad host range DNA cloning system for gram-negative bacteria: construction of a gene bank of Rhizobium meliloti. *Proc Natl Acad Sci U S A* **77**: 7347-7351

Feldman MF, Marolda CL, Monteiro MA, Perry MB, Parodi AJ, & Valvano MA (1999) The activity of a putative polyisoprenol-linked sugar translocase (Wzx) involved in Escherichia coli O antigen assembly is independent of the chemical structure of the O repeat. *J Biol Chem* **274**: 35129-35138

Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA, & Aebi M (2005) Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli. *Proc Natl Acad Sci U S A* **102**: 3016-3021

Feng L, Perepelov AV, Zhao G, Shevelev SD, Wang Q, Senchenkova SN, Shashkov AS, Geng Y, Reeves PR, Knirel YA, & Wang L (2007) Structural and genetic evidence that the Escherichia coli O148 O antigen is the precursor of the Shigella dysenteriae type 1 O antigen and identification of a glucosyltransferase gene. *Microbiology* **153**: 139-147

Foley SL, Lynne AM, & Nayak R (2009) Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infect Genet Evol* **9**: 430-440

Forsberg LS, Noel KD, Box J, & Carlson RW (2003) Genetic locus and structural characterization of the biochemical defect in the O-antigenic polysaccharide of the symbiotically deficient Rhizobium etli mutant, CE166. Replacement of N-acetylquinovosamine with its hexosyl-4-ulose precursor. *J Biol Chem* **278**: 51347-51359

Franco AV, Liu D, & Reeves PR (1996) A Wzz (Cld) protein determines the chain length of K lipopolysaccharide in Escherichia coli O8 and O9 strains. *J Bacteriol* **178**: 1903-1907

Fredriksson-Ahomaa M (2007) Yersinia enterocolitica and Yersinia pseudotuberculosis. In *Infectious Disease: Foodborne Diseases,* Simjee S (ed) pp 79-113. Humana Press Inc.: Totowa,NJ

Gronow S, Xia G, & Brade H (2010) Glycosyltransferases involved in the biosynthesis of the inner core region of different lipopolysaccharides. *Eur J Cell Biol* **89**: 3-10

Guo H, Yi W, Song JK, & Wang PG (2008) Current understanding on biosynthesis of microbial polysaccharides. *Curr Top Med Chem* **8**: 141-151

Guo L, Lim KB, Poduje CM, Daniel M, Gunn JS, Hackett M, & Miller SI (1998) Lipid A acylation and bacterial resistance against vertebrate antimicrobial peptides. *Cell* **95**: 189-198

Heinrichs DE, Yethon JA, Amor PA, & Whitfield C (1998a) The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of Escherichia coli. The R1 core-specific beta-glucosyltransferase provides a novel attachment site for O-polysaccharides. *J Biol Chem* **273**: 29497-29505

Heinrichs DE, Yethon JA, & Whitfield C (1998b) Molecular basis for structural diversity in the core regions of the lipopolysaccharides of Escherichia coli and Salmonella enterica. *Mol Microbiol* **30**: 221-232

Hong M & Payne SM (1997) Effect of mutations in Shigella flexneri chromosomal and plasmid-encoded lipopolysaccharide genes on invasion and serum resistance. *Mol Microbiol* **24**: 779-791

Kaluzny K, Abeyrathne PD, & Lam JS (2007) Coexistence of two distinct versions of O-antigen polymerase, Wzy-alpha and Wzy-beta, in Pseudomonas aeruginosa serogroup O2 and their contributions to cell surface diversity. *J Bacteriol* **189**: 4141-4152

Kanamori M (1976) Biological activities of endotoxins from Yersinia enterocolitica. *Jpn J Microbiol* **20**: 273-280

Kaniuk NA, Vinogradov E, & Whitfield C (2004) Investigation of the structural requirements in the lipopolysaccharide core acceptor for ligation of O antigens in the genus Salmonella: WaaL "ligase" is not the sole determinant of acceptor specificity. *J Biol Chem* **279**: 36470-36480

Kikuchi A, Elseviers D, & Gorini L (1975) Isolation and characterization of lambda transducing bacteriophages for argF, argI and adjacent genes. *J Bacteriol* **122**: 727-742

Kiljunen S, Hakala K, Pinta E, Huttunen S, Pluta P, Gador A, Lonnberg H, & Skurnik M (2005) Yersiniophage phiR1-37 is a tailed bacteriophage having a 270 kb DNA genome with thymidine replaced by deoxyuridine. *Microbiology* **151**: 4093-4102

Kim HM, Park BS, Kim JI, Kim SE, Lee J, Oh SC, Enkhbayar P, Matsushima N, Lee H, Yoo OJ, & Lee JO (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **130**: 906-917

Knirel I, Dentovskaya SV, Bystrova OV, Kocharova NA, Senchenkova SN, Shaikhutdinova RZ, Titareva GM, Bakhteeva IV, Lindner B, Pier GB, & Anisimov AP (2006) Relationship of structural variations in the lipopolysaccharide of Yersinia pestis to resistance to antimicrobial peptides and complement-mediated killing. *9th International Symposium on Yersinia, Program and Abstracts book S7:3*: 23-24

Knirel I & Kochetkov NK (1994) Structure of lipopolysaccharides from gram-negative bacteria III Structure of O-specific polysacchairdes. *Biokhimiia* **59,12**: 1784-1851

Knirel YA (2010**)** personal communication

Knirel YA (2009) O-Specific polysaccharides of Gram-negative bacteria. In *Microbial glycobiology: Structures, relevance and applications,* Moran A, Brennan P, Holst O, Von Itztein M (eds) pp 57-73. Elsevier

Lairson LL, Henrissat B, Davies GJ, & Withers SG (2008) Glycosyltransferases: structures, functions, and mechanisms. *Annu Rev Biochem* **77**: 521-555

Lan R & Reeves PR (2000) Intraspecies variation in bacterial genomes: the need for a species genome concept. *Trends Microbiol* **8**: 396-401

Larue K, Kimber MS, Ford R, & Whitfield C (2009) Biochemical and structural analysis of bacterial O-antigen chain length regulator proteins reveals a conserved quaternary structure. *J Biol Chem* **284**: 7395-7403

Lehrer J, Vigeant KA, Tatar LD, & Valvano MA (2007) Functional characterization and membrane topology of Escherichia coli WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial common antigen and O-antigen lipopolysaccharide. *J Bacteriol* **189**: 2618-2628

Marolda CL, Haggerty ER, Lung M, & Valvano MA (2008) Functional analysis of predicted coiled-coil regions in the Escherichia coli K-12 O-antigen polysaccharide chain length determinant Wzz. *J Bacteriol* **190**: 2128-2137

Marolda CL, Tatar LD, Alaimo C, Aebi M, & Valvano MA (2006) Interplay of the Wzx translocase and the corresponding polymerase and chain length regulator proteins in the translocation and periplasmic assembly of lipopolysaccharide o antigen. *J Bacteriol* **188**: 5124-5135

Marolda CL, Vicarioli J, & Valvano MA (2004) Wzx proteins involved in biosynthesis of O antigen function in association with the first sugar of the O-specific lipopolysaccharide subunit. *Microbiology* **150**: 4095-4105

Merhej V, Adekambi T, Pagnier I, Raoult D, & Drancourt M (2008) Yersinia massiliensis sp. nov., isolated from fresh water. *Int J Syst Evol Microbiol* **58**: 779-784

Metcalf WW, Jiang W, & Wanner BL (1994) Use of the rep technique for allele replacement to construct new Escherichia coli hosts for maintenance of R6K gamma origin plasmids at different copy numbers. *Gene* **138**: 1-7

Michiels T & Cornelis GR (1991) Secretion of hybrid proteins by the Yersinia Yop export system. *J Bacteriol* **173**: 1677-1685

Miller WL (2006) Investigation of enzymes involved in sugar-nucleotide biosynthesis pathways. *PhD Thesis University of Guelph*

Morona R, Purins L, Tocilj A, Matte A, & Cygler M (2009) Sequence-structure relationships in polysaccharide co-polymerase (PCP) proteins. *Trends Biochem Sci* **34**: 78-84

Morona R, Van Den Bosch L, & Daniels C (2000) Evaluation of Wzz/MPA1/MPA2 proteins based on the presence of coiled-coil regions. *Microbiology* **146 (Pt 1)**: 1-4

Morona R, van den Bosch L, & Manning PA (1995) Molecular, genetic, and topological characterization of O-antigen chain length regulation in Shigella flexneri. *J Bacteriol* **177**: 1059-1068

Murray GL, Attridge SR, & Morona R (2006) Altering the length of the lipopolysaccharide O antigen has an impact on the interaction of Salmonella enterica serovar Typhimurium with macrophages and complement. *J Bacteriol* **188**: 2735-2739

Murray GL, Attridge SR, & Morona R (2003) Regulation of Salmonella typhimurium lipopolysaccharide O antigen chain length is required for virulence; identification of FepE as a second Wzz. *Mol Microbiol* **47**: 1395-1406

Najdenski H, Golkocheva E, Vesselinova A, Bengoechea JA, & Skurnik M (2003) Proper expression of the O-antigen of lipopolysaccharide is essential for the virulence of Yersinia enterocolitica O:8 in experimental oral infection of rabbits. *FEMS Immunol Med Microbiol* **38**: 97-106

Nesper J, Lauriano CM, Klose KE, Kapfhammer D, Kraiss A, & Reidl J (2001) Characterization of Vibrio cholerae O1 El tor galU and galE mutants: influence on lipopolysaccharide structure, colonization, and biofilm formation. *Infect Immun* **69**: 435-445

Nikaido H (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* **67**: 593-656

Oertelt C, Lindner B, Skurnik M, & Holst O (2001) Isolation and structural characterization of an R-form lipopolysaccharide from Yersinia enterocolitica serotype O:8. *Eur J Biochem* **268**: 554-564

Pajunen M, Kiljunen S, & Skurnik M (2000) Bacteriophage phiYeO3-12, specific for Yersinia enterocolitica serotype O:3, is related to coliphages T3 and T7. *J Bacteriol* **182**: 5114-5120

Paulsen IT, Beness AM, & Saier MH,Jr (1997) Computer-based analyses of the protein constituents of transport systems catalysing export of complex carbohydrates in bacteria. *Microbiology* **143 (Pt 8)**: 2685-2699

Pekkola-Heino K, Viljanen MK, Stahlberg TH, Granfors K, & Toivanen A (1987) Monoclonal antibodies reacting selectively with core and O-polysaccharide

of Yersinia enterocolitica O:3 lipopolysaccharide. *Acta Pathol Microbiol Immunol Scand C* **95**: 27-34

Perez JM, McGarry MA, Marolda CL, & Valvano MA (2008) Functional analysis of the large periplasmic loop of the Escherichia coli K-12 WaaL O-antigen ligase. *Mol Microbiol* **70**: 1424-1440

Perez-Gutierrez C, Llobet E, Llompart CM, Reines M, & Bengoechea JA (2010) Role of lipid A acylation on Yersinia enterocolitica virulence. *Infect Immun* **78**: 2768-2781

Perry RD & Fetherston JD (1997) Yersinia pestis--etiologic agent of plague. *Clin Microbiol Rev* **10**: 35-66

Radziejewska-Lebrecht J, Shashkov AS, Stroobant V, Wartenberg K, Warth C, & Mayer H (1994) The inner core region of Yersinia enterocolitica Ye75R (0:3) lipopolysaccharide. *Eur J Biochem* **221**: 343-351

Radziejewska-Lebrecht J, Skurnik M, Shashkov AS, Brade L, Rozalski A, Bartodziejska B, & Mayer H (1998) Immunochemical studies on R mutants of Yersinia enterocolitica O:3. *Acta Biochim Pol* **45**: 1011-1019

Raetz CR, Guan Z, Ingram BO, Six DA, Song F, Wang X, & Zhao J (2009) Discovery of new biosynthetic pathways: the lipid A story. *J Lipid Res* **50 Suppl**: S103-8

Raetz CR, Reynolds CM, Trent MS, & Bishop RE (2007) Lipid A modification systems in gram-negative bacteria. *Annu Rev Biochem* **76**: 295-329

Raetz CR & Whitfield C (2002) Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**: 635-700

Reeves P (1993) Evolution of Salmonella O antigen variation by interspecific gene transfer on a large scale. *Trends Genet* **9**: 17-22

Rick PD, Hubbard GL, Kitaoka M, Nagaki H, Kinoshita T, Dowd S, Simplaceanu V, & Ho C (1998) Characterization of the lipid-carrier involved in the synthesis of enterobacterial common antigen (ECA) and identification of a novel phosphoglyceride in a mutant of Salmonella typhimurium defective in ECA synthesis. *Glycobiology* **8**: 557-567

Rietschel ET, Kirikae T, Schade FU, Mamat U, Schmidt G, Loppnow H, Ulmer AJ, Zahringer U, Seydel U, & Di Padova F (1994) Bacterial endotoxin: molecular relationships of structure to activity and function. *FASEB J* **8**: 217-225

Robbins PW, Bray D, Dankert BM, & Wright A (1967) Direction of chain growth in polysaccharide synthesis. *Science* **158**: 1536-1542

Rocchetta HL, Burrows LL, & Lam JS (1999) Genetics of O-antigen biosynthesis in Pseudomonas aeruginosa. *Microbiol Mol Biol Rev* **63**: 523-553

Ruiz N, Gronenberg LS, Kahne D, & Silhavy TJ (2008) Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of Escherichia coli. *Proc Natl Acad Sci U S A* **105**: 5537-5542

Rush JS, Alaimo C, Robbiani R, Wacker M, & Waechter CJ (2010) A novel epimerase that converts GlcNAc-P-P-undecaprenol to GalNAc-P-P-undecaprenol in Escherichia coli O157. *J Biol Chem* **285**: 1671-1680

Sambrook J & Russel D (2001) *Molecular Cloning, A Laboratory Manual.* Cold Sprong Harbor Laboratory Press: New York

Samuel G & Reeves P (2003) Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. *Carbohydr Res* **338**: 2503-2519

Shepherd JG, Wang L, & Reeves PR (2000) Comparison of O-antigen gene clusters of Escherichia coli (Shigella) sonnei and Plesiomonas shigelloides O17: sonnei gained its current plasmid-borne O-antigen genes from P. shigelloides in a recent event. *Infect Immun* **68**: 6056-6061

Shibaev V (1986) Biosynthesis of bacterial polysaccharide chains composed of repeating units. In *Advances in carbohydrate chemistry and biochemistry,* Anonymous pp 277-339. Academic Press

Sirisena DM & Skurnik M (2003) Mutations in the genes for synthesis of the outer core region of the lipopolysaccharide of Yersinia enterocolitica O:3. *J Appl Microbiol* **94**: 686-692

Skurnik M (1984) Lack of correlation between the presence of plasmids and fimbriae in Yersinia enterocolitica and Yersinia pseudotuberculosis. *J Appl Bacteriol* **56**: 355-363

Skurnik M, Peippo A, & Ervela E (2000) Characterization of the O-antigen gene clusters of Yersinia pseudotuberculosis and the cryptic O-antigen gene cluster of Yersinia pestis shows that the plague bacillus is most closely related to and has evolved from Y. pseudotuberculosis serotype O:1b. *Mol Microbiol* **37**: 316-330

Skurnik M & Toivanen P (1993) Yersinia enterocolitica lipopolysaccharide: genetics and virulence. *Trends Microbiol* **1**: 148-152

Skurnik M, Venho R, Bengoechea JA, & Moriyon I (1999) The lipopolysaccharide outer core of Yersinia enterocolitica serotype O:3 is required for virulence and plays a role in outer membrane integrity. *Mol Microbiol* **31**: 1443-1462

Skurnik M, Venho R, Toivanen P, & al-Hendy A (1995) A novel locus of Yersinia enterocolitica serotype O:3 involved in lipopolysaccharide outer core biosynthesis. *Mol Microbiol* **17**: 575-594

Sperandeo P, Cescutti R, Villa R, Di Benedetto C, Candia D, Deho G, & Polissi A (2007) Characterization of lptA and lptB, two essential genes implicated in

lipopolysaccharide transport to the outer membrane of Escherichia coli. *J Bacteriol* **189**: 244-253

Sperandeo P, Lau FK, Carpentieri A, De Castro C, Molinaro A, Deho G, Silhavy TJ, & Polissi A (2008) Functional analysis of the protein machinery required for transport of lipopolysaccharide to the outer membrane of Escherichia coli. *J Bacteriol* **190**: 4460-4469

Sprague LD & Neubauer H (2005) Yersinia aleksiciae sp. nov. *Int J Syst Evol Microbiol* **55**: 831-835

Sprague LD, Scholz HC, Amann S, Busse HJ, & Neubauer H (2008) Yersinia similis sp. nov. *Int J Syst Evol Microbiol* **58**: 952-958

Stenutz R, Weintraub A, & Widmalm G (2006) The structures of Escherichia coli O-polysaccharide antigens. *FEMS Microbiol Rev* **30**: 382-403

Stevenson G, Lan R, & Reeves PR (2000) The colanic acid gene cluster of Salmonella enterica has a complex history. *FEMS Microbiol Lett* **191**: 11-16

Strauch E, Kaspar H, Schaudinn C, Damasko C, Konietzny A, Dersch P, Skurnik M, & Appel B (2003) Analysis of enterocoliticin, a phage tail-like bacteriocin. *Adv Exp Med Biol* **529**: 249-251

Strauch E, Kaspar H, Schaudinn C, Dersch P, Madela K, Gewinner C, Hertwig S, Wecke J, & Appel B (2001) Characterization of enterocoliticin, a phage tail-like bacteriocin, and its effect on pathogenic Yersinia enterocolitica strains. *Appl Environ Microbiol* **67**: 5634-5642

Sugiyama T, Kido N, Kato Y, Koide N, Yoshida T, & Yokochi T (1998) Generation of Escherichia coli O9a serotype, a subtype of E. coli O9, by transfer of the wb* gene cluster of Klebsiella O3 into E. coli via recombination. *J Bacteriol* **180**: 2775-2778

Tatar LD, Marolda CL, Polischuk AN, van Leeuwen D, & Valvano MA (2007) An Escherichia coli undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling. *Microbiology* **153**: 2518-2529

Thibodeaux CJ, Melancon CE, & Liu HW (2007) Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**: 1008-1016

Thoerner P, Bin Kingombe CI, Bogli-Stuber K, Bissig-Choisat B, Wassenaar TM, Frey J, & Jemmi T (2003) PCR detection of virulence genes in Yersinia enterocolitica and Yersinia pseudotuberculosis and investigation of virulence gene distribution. *Appl Environ Microbiol* **69**: 1810-1816

Tocilj A, Munger C, Proteau A, Morona R, Purins L, Ajamian E, Wagner J, Papadopoulos M, Van Den Bosch L, Rubinstein JL, Fethiere J, Matte A, & Cygler M (2008) Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat Struct Mol Biol* **15**: 130-138

Tomshich SV, Gorshkova RP, & Ovodoy YS (1987) Structural investigation of the lipopolysaccharide of Yersinia enterocolitica serovar 0:8. *J Chem Nat Comp* **23**: 545-551

Touze T, Tran AX, Hankins JV, Mengin-Lecreulx D, & Trent MS (2008) Periplasmic phosphorylation of lipid A is linked to the synthesis of undecaprenyl phosphate. *Mol Microbiol* **67**: 264-277

Valvano MA (2008) Undecaprenyl phosphate recycling comes out of age. *Mol Microbiol* **67**: 232-235

Valvano MA (2003) Export of O-specific lipopolysaccharide. *Front Biosci* **8**: s452-71

Wang X & Quinn PJ (2010) Lipopolysaccharide: Biosynthetic pathway and structure modification. *Prog Lipid Res* **49**: 97-107

Whitfield C (1995) Biosynthesis of lipopolysaccharide O antigens. *Trends Microbiol* **3**: 178-185

Whitfield C, Amor PA, & Koplin R (1997) Modulation of the surface architecture of gram-negative bacteria by the action of surface polymer:lipid A-core ligase and by determinants of polymer chain length. *Mol Microbiol* **23**: 629-638

Whitfield C & Roberts IS (1999) Structure, assembly and regulation of expression of capsules in Escherichia coli. *Mol Microbiol* **31**: 1307-1319

Wilson KJ, Sessitsch A, Corbo JC, Giller KE, Akkermans AD, & Jefferson RA (1995) beta-Glucuronidase (GUS) transposons for ecological and genetic studies of rhizobia and other gram-negative bacteria. *Microbiology* **141 (Pt 7)**: 1691-1705

Woodcock DM, Crowther PJ, Doherty J, Jefferson S, DeCruz E, Noyer-Weidner M, Smith SS, Michael MZ, & Graham MW (1989) Quantitative evaluation of Escherichia coli host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res* **17**: 3469-3478

Woodward R, Yi W, Li L, Zhao G, Eguchi H, Sridhar PR, Guo H, Song JK, Motari E, Cai L, Kelleher P, Liu X, Han W, Zhang W, Ding Y, Li M, & Wang PG (2010) In vitro bacterial polysaccharide biosynthesis: defining the functions of Wzy and Wzz. *Nat Chem Biol* **6**: 418-423

Wu T, McCandlish AC, Gronenberg LS, Chng SS, Silhavy TJ, & 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

Yanisch-Perron C, Vieira J, & Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119

Zahringer U, Lindner B, & Rietschel ET (1994) Molecular structure of lipid A, the endotoxic center of bacterial lipopolysaccharides. *Adv Carbohydr Chem Biochem* **50**: 211-276

Zhang L, Radziejewska-Lebrecht J, Krajewska-Pietrasik D, Toivanen P, & Skurnik M (1997) Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of Yersinia enterocolitica serotype O:8. *Mol Microbiol* **23**: 63-76

Zhang L & Skurnik M (1994) Isolation of an R- M+ mutant of Yersinia enterocolitica serotype O:8 and its application in construction of rough mutants utilizing mini-Tn5 derivatives and lipopolysaccharide-specific phage. *J Bacteriol* **176**: 1756-1760

Zhang L, Toivanen P, & Skurnik M (1996) The gene cluster directing O-antigen biosynthesis in Yersinia enterocolitica serotype 0:8: identification of the genes for mannose and galactose biosynthesis and the gene for the O-antigen polymerase. *Microbiology* **142 (Pt 2)**: 277-288

Zhao X, Creuzenet C, Belanger M, Egbosimba E, Li J, & Lam JS (2000) WbpO, a UDP-N-acetyl-D-galactosamine dehydrogenase from Pseudomonas aeruginosa serotype O6. *J Biol Chem* **275**: 33252-33259

Zhou D, Tong Z, Song Y, Han Y, Pei D, Pang X, Zhai J, Li M, Cui B, Qi Z, Jin L, Dai R, Du Z, Wang J, Guo Z, Wang J, Huang P, & Yang R (2004) Genetics of metabolic variations between Yersinia pestis biovars and the proposal of a new biovar, microtus. *J Bacteriol* **186**: 5147-5152