Salivary scavenger and agglutinin
SALSA
in innate immunity

Martin Parnov Reichhardt
Immunobiology Research Program
Department of Bacteriology and Immunology
Faculty of Medicine
University of Helsinki
Finland

Academic dissertation
To be publicly discussed with the permission of the Medical Faculty, University of Helsinki, in lecture hall 1, Haartmaninkatu 3, Helsinki, on the 7th August 2015 at 12 o'clock noon
Cover photo:
SALSA and fibronectin
co-localization in human placenta.
Photo by the author.

Dissertationes Scholae Doctoralis Ad Sanitatem Investigandum
Universitatis Helsinkiensis
ISSN 2342-3161 (print)
ISSN 2342-317X (online)

Printed at Hansaprint Oy
Vantaa, Finland 2015
ISBN 978-951-51-1390-0 (paperback)
ISBN 978-951-51-1391-7 (PDF)
http://ethesis.helsinki.fi
© 2015 Martin Parnov Reichhardt
Supervisors

Professor Seppo Meri, M.D, Ph.D.
Immunobiology Research Program and Department of
Bacteriology and Immunology, Faculty of Medicine,
University of Helsinki, Finland

and

Docent Hanna Jarva, M.D, Ph.D.
Immunobiology Research Program and Department of
Bacteriology and Immunology, Faculty of Medicine,
University of Helsinki, Finland

Reviewers

Professor Olli Vainio, M.D, Ph.D.
Research group of Biomedicine, Department of Medical
Microbiology and Immunology, University of Oulu, Finland
and Nordlab Oulu, Finland

and

Docent Anna-Maija Haapala, M.D, Ph.D.
Fimlab Laboratories, Tampere and University of Turku,
Finland

Opponent

Professor Roland Jonsson, D.M.D., Ph.D.
Broegelmann Research Laboratory, Department of Clinical
Science, University of Bergen, Bergen, Norway and
Department of Rheumatology, Haukeland University
Hospital, Bergen, Norway.
To mum,
for always encouraging me
to pursue ALL of my dreams
If happy little bluebirds fly
beyond the rainbow why, oh why
can’t I?

*Judy Garland, 1939*
Acknowledgements

This thesis work was performed as part of the Research Programs Unit in the Immunobiology Research Program at the Department of Bacteriology and Immunology in the Haartman Institute, at the University of Helsinki, Finland. I thank Professor Mikael Skurnik for supporting the exchange of knowledge between individuals as the head of the Immunobiology Research Program.

The project was financially supported by the Doctoral Programme in Biomedicine under the Doctoral School in Health Sciences at the University of Helsinki. Further funding was provided by the Federation of European Microbiological Societies, the Centre for International Mobility, the Scandinavian Society of Immunology, the Orion foundation, the Sigrid Juselius foundation, the Helsinki University Hospital funds (EVO) and the Academy of Finland.

I am honored that Professor Roland Jonsson accepted the invitation to be my opponent. I greatly appreciate the work of my reviewers, Professor Olli Vainio and Docent Anna-Maija Haapala, and I thank them for their contributions to the improvement of my final dissertation. In addition, I thank Sampsa Mattikainen and Marc Baumann for being members of my thesis committee and guiding the progress of my doctoral studies.
My deepest gratitude I owe to my supervisor Professor Seppo Meri for trusting in my science, although it was only one of my two big passions. I thank you for the opportunity to do my research, for teaching me how to see the small molecular interactions in a broader biological perspective and for allowing me great independence to follow my ideas. Laboratory work is a craftsmanship, and I could have asked for no greater teacher than my supervisor Hanna Jarva. I thank her deeply for showing me the art of precision and for always being there to help me.

I wish to express my deepest gratitude to my co-authors Vuokko Loimaranta, Steffen Thiel, Jukka Finne, Mark de Been, Juan Miguel Rodriguez, Esther Jimenez Quintana, Willem de Vos, Anna Inkeri Lokki, Hannele Laivuori, Piia Vuorela, Andreas Glasner, Monika Siwetz and Berthold Huppertz. Without your contributions of sample collection and preparation, practical work in the laboratory, and theoretical knowledge, the exploration of the SALSA molecule would not have moved forward. I am indebted to you.

To my colleagues in the Meri lab I will forever be grateful for your technical assistance, our theoretical discussions and above all your patience towards the dancing scientist. I especially wish to thank Marcel Messing for invaluable help with every aspect of laboratory life and Tobias Freitag for broadening my mind. Thank you Judith Klievink and Anna-Helena Saariaho for your friendship and for adding sunshine
to everyday life at the University of Helsinki. The happy-pills will not be forgotten!

I wish to thank my colleagues at the Haartman Institute, especially Sakari Jokiranta, Hanne Amdahl, Satu Hyvärinen and Taru Meri. In addition, a genuine thank you goes to my office mates, Dawitt Yohannes, Nelli Heikkilä, Iivo Hetemäki, Rigbe Weldatsadik and Mabruka Salem for always providing a warm and positive work environment. I appreciate the help I have received from Kirsi Udueze, Taija Pietilä and Liisa Penttilä in the office. I offer sincere gratitude to Heidi Sillanpää for suffering with me through the Immunobiology exam and for invaluable help with preparing my thesis.

A heartfelt thank you to Maija Salminen for constantly teaching me about teamwork, dedication and even a bit of magic. Thank you Harri Antikainen for teaching me about the whole human being.

Finally I wish to thank my mother, Birgitte, my father, Poul Erik, my family and my love, John, for your infinite care and your support of my work, even though it was accomplished in a place far away from home.

A fool remains only he, who does not dare to ask questions.

Martin Parnov Reichhardt
Helsinki, August 2015
# Table of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>9</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>13</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>17</td>
</tr>
<tr>
<td>LIST OF ORIGINAL PUBLICATIONS</td>
<td>20</td>
</tr>
<tr>
<td>ABBREVIATIONS</td>
<td>21</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>24</td>
</tr>
<tr>
<td>LITERATURE REVIEW</td>
<td>27</td>
</tr>
<tr>
<td>MUCOSAL SURFACES</td>
<td>27</td>
</tr>
<tr>
<td>PHYSICAL STRUCTURE</td>
<td>27</td>
</tr>
<tr>
<td>CELLS OF THE MUCOSAL SURFACES</td>
<td>27</td>
</tr>
<tr>
<td>BACTERIAL COLONIZATION OF THE MUCOSA</td>
<td>28</td>
</tr>
<tr>
<td>IMMUNOLOGY OF THE MUCOSAL SURFACE</td>
<td>32</td>
</tr>
<tr>
<td>WHEN THE BARRIER IS BREACHED – ENCOUNTER WITH THE</td>
<td>34</td>
</tr>
<tr>
<td>COMPLEMENT SYSTEM</td>
<td>34</td>
</tr>
<tr>
<td>COMPLEMENT ACTIVATION</td>
<td>34</td>
</tr>
<tr>
<td>COMPLEMENT REGULATION</td>
<td>38</td>
</tr>
<tr>
<td>COMPLEMENT RECEPTORS</td>
<td>39</td>
</tr>
<tr>
<td>FUNCTIONS OF COMPLEMENT</td>
<td>40</td>
</tr>
<tr>
<td>COMPLEMENT AND MUCOSAL SURFACES</td>
<td>42</td>
</tr>
</tbody>
</table>
Table of contents

**PLACENTA**

- Placental Anatomy and Development 43
- Immunology of the Placenta 45
- Complement in the Placenta 46

**SALSA PROTEIN**

- Discovery 47
- DMBT1 Gene 48
- Glycosylation of the SALSA Protein 49
- SALSA Across Species 51
- SALSA Expression in Humans 51

**FUNCTIONS OF SALSA**

- SALSA and Epithelial Differentiation 56
- Functions of SALSA in the Innate Immune System 57

**AIMS OF THE STUDY** 65

**MATERIALS AND METHODS** 66

**HUMAN SAMPLES**

- Amniotic Fluid 67
- Intestinal Samples 67
- Placental Samples 69

**PROTEIN LEVEL MEASUREMENTS** 69

- Quantification of SALSA in Amniotic Fluid by ELISA 69
- LC-MS/MS Mass Spectrometry 72

**PROTEIN VISUALIZATION ASSAYS** 73

- Western Blotting 73
- Immunohistochemistry 73

**PROTEIN INTERACTION ASSAYS** 74

- ELISA Binding Assays 74
- ELISA Competition Assays 75
# Table of contents

## Complement assays

- Measurement of complement activation by SALSA in solution 76
- Measurement of complement activation by surface-coated SALSA 76
- Effect of SALSA on complement activation by *C. albicans* 76

## Bacterial binding assays

- SALSA-mediated inhibition of MBL binding to microorganisms 78

## Coagulation assays

- Effect of soluble SALSA on coagulation 79
- Coagulation in the presence of surface-coated SALSA 79

## Statistical analysis

80

## Results and discussion

81

## SALSA in body fluids

- SALSA in amniotic fluid (II, III) 82
- Amniotic fluid SALSA levels in normal and complicated pregnancies (II, III) 84
- SALSA in the infant intestine (II) 88
- Comparison of SALSA in amniotic fluid, meconium and feces (II) 90

## Ligand binding by SALSA

- Novel ligands of SALSA (I, II, III) 94
- SALSA-MBL interaction (I) 96

## Complement regulation by SALSA

- Complement activation by surface-bound SALSA (I) 98
- Complement activation by fluid-phase SALSA (I) 99
Table of contents

**BINDING OF SALSA TO BACTERIA** 101
- Bacterial binding by purified SALSA 101
- Bacterial binding by SALSA in biological fluids (II) 104

**SALSA AT THE FETO-MATERNAL INTERFACE** 106
- SALSA localization in placenta (III) 106
- SALSA in fibrinoids (III) 108
- SALSA in 1st trimester placenta 110
- Effect of SALSA on blood clotting (III) 111
- Co-localization of SALSA with complement and fibronectin (III) 113

**CONCLUSION** 116

**REFERENCES** 120
Abstract

To live a healthy life, humans need to co-exist with foreign organisms. These consist of the thousands of different types of microbes that colonize the human body. But also, in the case of a pregnant woman, the fetus can be viewed as a foreign organism. To avoid disease, the barriers of the human body, e.g. the mucosal surfaces, must be maintained. Here the innate immune defense system plays an important role.

The salivary scavenger and agglutinin (SALSA), also known as gp340, DMBT1 and SAG, is a molecule found at most mucosal surfaces. SALSA is associated with the epithelium or secreted into the lining fluids, such as tears, saliva and mucus in the respiratory tract. SALSA is known to bind and agglutinate a broad spectrum of bacteria, as well as viruses, and thus play a role in the innate immune defense against invading microbes. The effect of SALSA is mediated in concert with several other defense molecules such as IgA, surfactant proteins A and D, and the complement component C1q. These have all been shown to be ligands of SALSA. Alongside the role of SALSA in innate immunity, evidence for a function in epithelial and stem cell differentiation has emerged.

This thesis work has addressed the function of SALSA in innate immunity, especially in early life. SALSA was found in
Abstract

the amniotic fluid and in meconium and feces of newborns. In fact, SALSA was among the most abundant proteins in the intestines of newborn children. By comparing the SALSA protein in the different samples we found size polymorphisms, varying from one individual to another, but also from compartment to compartment within the same individual. Specifically, we found structural variations in SALSA correlating to the known bacterial binding peptide sequence, SRCRP2, and the putative polymerization domain, the zona pellucida domain. These differences were found to alter the ability of SALSA to bind known endogenous and bacterial ligands.

SALSA was also found to be expressed in the human placental and decidual tissues. In the 1st trimester of pregnancy, SALSA was detected sporadically in maternal decidual capillaries. Closer to term SALSA was found to be expressed by the syncytiotrophoblast layer of the placental villous trees. In certain sites, e.g. at disrupted and damaged areas of the syncytium, SALSA was found deposited into fibrinoid formations. It partially co-localized with the fibrinoid component fibronectin. Damage of the syncytiotrophoblast layer is a common histological finding of several pregnancy complications. We thus investigated the presence of SALSA in amniotic fluids and/or placentas from patients with pre-eclampsia, intra-uterine growth restriction, diabetes mellitus type 1 and gestational diabetes. SALSA levels were increased in amniotic fluid samples collected before the 20th week of gestation from women who
later developed pre-eclampsia, but no other differences between the groups were observed.

Complement activation has been observed at the feto-maternal interface of both healthy and complicated pregnancies. SALSA had previously been found to bind C1q. Thus, it was of interest to investigate the ability of SALSA to interact directly with the complement system. We found that SALSA bound to both mannan-binding lectin and to some extent to all three ficolins (H, L and M). SALSA activated complement, when it was bound to a surface. In contrast, fluid-phase SALSA was able to inhibit the deposition of complement on SALSA non-binding microbial surfaces. It thus acted in dual fashion to target complement attack.

In the human placenta we observed C1q-targeting of the SALSA-positive fibrinoid formations. C1q and complement are known to function in the clearance of apoptotic cells and debris. Thereby SALSA and complement probably have a cooperative function in the containment and clearance of the injured structures, thus linking its innate immune activity with the maintenance of tissue homeostasis.
List of original publications

This thesis is based on the following original articles, which are referred to by their Roman numerals:


The original publications have been reprinted with the kind permission of the copyright holders.
Abbreviations

AF  Amniotic fluid
AP  Alkaline phosphatase
APC  Antigen-presenting cell
BSA  Bovine serum albumin
C  Complement
CR  Complement receptor
CRD  Carbohydrate recognition domain
CRIg  Complement receptor of the immunoglobulin family
CRP  C-reactive protein
CUB  C1r/C1s, urchin embryonic growth factor and bone morphogenetic protein-1
C3aR  C3a receptor
C5aR  C5a receptor
DM  Diabetes mellitus type 1
DMBT1  Deleted in malignant brain tumors 1
DSS  Dextran sulfate sodium
EDTA  Ethylene diamine tetraacetic acid
EGTA  Ethylene glycol tetraacetic acid
ELISA  Enzyme linked immuno-sorbent assay
Fv  Fetal vessels
GAS  Group A streptococcus, *S. pyogenes*
GBS  Group B streptococcus, *S. agalactiae*
GDM  Gestational diabetes
GlcNAc  N-acetylglucosamine
gp340  Glycoprotein of 340 kDa
HIS  Heat-inactivated serum
HIV-1  Human immunodeficiency virus type 1
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HRP</td>
<td>Horseradish peroxidase</td>
</tr>
<tr>
<td>IAV</td>
<td>Influenza A virus</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IUGR</td>
<td>Intra-uterine growth restriction</td>
</tr>
<tr>
<td>IVS</td>
<td>Intervillous space</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>Le&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Lewis antigen a</td>
</tr>
<tr>
<td>Le&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Lewis antigen b</td>
</tr>
<tr>
<td>Le&lt;sup&gt;x&lt;/sup&gt;</td>
<td>Lewis antigen x</td>
</tr>
<tr>
<td>Le&lt;sup&gt;y&lt;/sup&gt;</td>
<td>Lewis antigen y</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>MAC</td>
<td>Membrane attack complex</td>
</tr>
<tr>
<td>MASP</td>
<td>MBL-associated serine protease</td>
</tr>
<tr>
<td>MBL</td>
<td>Mannose binding lectin</td>
</tr>
<tr>
<td>MFI</td>
<td>Mean fluorescence intensity</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>NB</td>
<td>Northern blotting</td>
</tr>
<tr>
<td>NLR</td>
<td>NOD-like receptor</td>
</tr>
<tr>
<td>NOD2</td>
<td>Nucleotide-binding oligomerization domain 2</td>
</tr>
<tr>
<td>NHS</td>
<td>Normal human serum</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>O/N</td>
<td>Over night</td>
</tr>
<tr>
<td>OPD</td>
<td>1,2-phenylenediamine</td>
</tr>
<tr>
<td>pAb</td>
<td>Polyclonal antibody</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PE</td>
<td>Pre-eclampsia</td>
</tr>
<tr>
<td>PRM</td>
<td>Pattern recognition molecule</td>
</tr>
<tr>
<td>PMA</td>
<td>Phorbol 12-myristate 13-acetate</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>rMASP-2</td>
<td>Recombinant MASP-2</td>
</tr>
</tbody>
</table>
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>rMBL</td>
<td>Recombinant MBL</td>
</tr>
<tr>
<td>rM-ficolin</td>
<td>Recombinant M-ficolin</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal ribonucleic acid</td>
</tr>
<tr>
<td>rSALSA</td>
<td>Recombinant SALSA</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcriptase polymerase chain reaction</td>
</tr>
<tr>
<td>SA</td>
<td>Spiral arteries</td>
</tr>
<tr>
<td>SAG</td>
<td>Salivary agglutinin</td>
</tr>
<tr>
<td>SALSA</td>
<td>Salivary scavenger and agglutinin</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>Se(+)</td>
<td>Secretor</td>
</tr>
<tr>
<td>Se(-)</td>
<td>Non-secretor</td>
</tr>
<tr>
<td>SID</td>
<td>SRCR interspersed domain</td>
</tr>
<tr>
<td>SpA</td>
<td>Surfactant protein A</td>
</tr>
<tr>
<td>SpD</td>
<td>Surfactant protein D</td>
</tr>
<tr>
<td>SRCR</td>
<td>Scavenger receptor cysteine-rich</td>
</tr>
<tr>
<td>SRCR-SF</td>
<td>Scavenger receptor cysteine-rich superfamily</td>
</tr>
<tr>
<td>SRCRP2</td>
<td>Scavenger receptor cysteine-rich peptide 2</td>
</tr>
<tr>
<td>STP</td>
<td>Serine threonine proline</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffered saline</td>
</tr>
<tr>
<td>TTSB</td>
<td>Tris saline buffer containing tween</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumor necrosis factor α</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
<tr>
<td>ZP</td>
<td>Zona pellucida</td>
</tr>
<tr>
<td>VBS</td>
<td>Veronal-buffered saline</td>
</tr>
<tr>
<td>WB</td>
<td>Western blotting</td>
</tr>
<tr>
<td>w/w</td>
<td>weight versus weight</td>
</tr>
</tbody>
</table>
Introduction

An individual human being is a unique organism distinct from all other living organisms. In pregnancy, however, we may think, that a fetus inside a pregnant woman is part of her. But in reality, within the mother there is a separate compartment for the child, and a barrier between the mother’s and the child’s tissues. In a smaller scale we know that the human body harbors thousands of different species of bacteria. In fact, on the average there are roughly 10 times as many bacterial cells living in the human body, as there are human cells. However, there is still a clear barrier keeping the bacteria separated from the parenteral human tissues. As such, we need to think of the surface of the human body not only as the skin and other visible parts, but also the surfaces within the human body, the mucosal surfaces, such as the gastro-intestinal, respiratory, urinary and the genital tracts and during pregnancy even the placenta.

Maintaining these barriers is essential for a healthy life. When the barriers are breached we become susceptible to diseases such as infections and, for pregnant women, to certain pregnancy complications. To prevent barrier damage, mechanical defenses exist, for example the gastric acid, the impenetrable skin, and the activity of cilia to constantly remove bacteria and waste products out of the lungs. However, for an efficient and more sophisticated protection the immune system is needed. It plays an
indispensable role in constantly monitoring the surfaces of the human body.

The immune defense system consists of an innate and an adaptive part. The innate immune system utilizes a number of evolutionarily conserved anti-microbial molecules as well as cells with receptors specialized in recognizing a vast array of microbial structures. In contrast, the adaptive immune system is only engaged upon specific challenge and presentation of an antigen. Activation of T-cells and B-cells of adaptive immunity leads to a more precise and long-lasting defense. T-cells direct cell-mediated immunity and B-cells develop into antibody-producing plasma cells. Because a newborn child has not yet encountered a broad spectrum of microbes, its adaptive immune system is still immature. Although some antibodies are transferred directly from mother to child via the placenta and breast milk, the child still relies mostly on the innate immune system in early life.

It is obvious that maternal immune activation against the “foreign” fetus would be catastrophic. In addition, human health relies on peaceful and synergistic interactions with the colonizing microbiota. Therefore, a strict regulation targeting the immune system towards some, but not all foreign organisms exists. When the monitoring cells of the innate immune system, such as dendritic cells and macrophages, meet a foreign organism their response and interaction with T-cells and adaptive immunity depends on the local environment of danger signals. Binding and activation of anti-microbial proteins and enzymatic cascades,
Introduction

such as the complement (C) system, are important features when the immune system decides whether to mount a full attack or simply tolerate the target. Therefore a malfunction of these systems may lead to unwanted immunological responses and subsequent diseases. Elucidating the role of these molecules is of utmost importance, and therefore the overall scope of this study. By understanding the basic physiological mechanisms we may also realize what goes wrong in the case of illness. Finally, understanding the basic pathophysiological mechanisms will help us finding new ways to cure the respective diseases – maybe, one day, with a pill of SALSA!
Mucosal surfaces

Physical structure
The barrier between the human body and the surrounding environment consists of the skin and the mucosal surfaces. The main mucosal surfaces are the mouth, the respiratory tract, the gastric sac and the intestines with variation observed between duodenum, jejunum, ileum and colon. While the skin makes up approximately 2 m$^2$, the mucosal surfaces cover up to 300 m$^2$, making them, by far, the largest interface between the human host and foreign organisms [197]. The skin is covered by several layers of dead and living epithelial cells providing an extensive mechanical barrier. However, the mucosal surfaces are only covered by a single layer of epithelial cells making the requirement for strong immunological regulation evident [23].

Cells of the mucosal surfaces
The key players at the mucosal surfaces are the single-layered epithelial cells. Seeded on a basement membrane these cells make up the main barrier to the environment. The polarized epithelial cells are covered with a thick layer of mucus on the luminal side designed to help them in the interactions with the colonizing bacteria [32]. The epithelial cells are mainly involved in absorption and digestion of
Literature review

nutrients, however, they have also been shown have very important immunological functions, e.g. the expression of specific microbial receptors \[^{23}\]. In addition to the epithelial cells, other cell types play important roles in maintaining the mucosal barrier such as goblet cells, endocrine cells and Paneth cells. These cells secrete a large number of substances involved in the interactions with the microbiota, e.g. mucus components, acid in the stomach, epithelial growth factors and antibacterial peptides \[^{197}\].

A great number of immunological cells are found both below the single layered epithelium and interspersed between the epithelial cells. Antigen-presenting cells (APCs) such as dendritic cells and specialized M-cells are found with direct contact to the gut lumen. In the underlying lamina propria both B-cells and T-cells gather in specialized compartments known as Peyer's patches. In addition, isolated lymphocytes and innate immune effector cells such as macrophages, natural killer cells and mast cells are found spread out in the entire subepithelial compartment \[^{22}\]. An overview is given in Figure 1.

**Bacterial colonization of the mucosa**

Humans are born virtually sterile, but immediately after birth the body is colonized by a multitude of microorganisms \[^{96}\]. Some studies have shown bacterial colonization of both amniotic fluid and infant meconium from healthy individuals, suggesting that bacteria may be present in the amniotic cavity already during pregnancy \[^{9, 77}\]. However, the
Figure 1: Immunology of the mucosal surfaces exemplified by the gut.

Epithelial cells line the surface of the gut, with dendritic cells protruding through the cell layer to monitor the lumen. M-cells are responsible for transporting luminal antigens to the structured lymphoid organs, Peyer's patches, with distinct T-cell areas (blue) and B-cell follicles (yellow). In addition, intra-epithelial T-cells are found scattered throughout the mucosal surface. When the epithelium is damaged, lumen defense molecules meet and interact with the cells and molecules from the tissue to protect against infection and to initiate healing.
main colonizing microbes appear from the surrounding environment, and in particular from the vaginal flora of the mother. Eventually the commensal flora of the human body displays a profound diversity with more than 1000 different species co-existing within the human host\textsuperscript{197}. Most of these bacteria exist in a symbiotic relationship with the human host (mutualism). However, disruption of this mutualistic balance can lead to disease. Under certain conditions bacteria may become opportunistic pathogens leading to a harmful infection of the host\textsuperscript{36}. In contrast, an over-reactive immune system may cause chronic inflammation such as in Chron's disease\textsuperscript{67,140}.

The specific mucosal tissues make up specific microenvironments, and therefore also attract the colonization of certain types of microorganisms\textsuperscript{116}. An example is the acid-tolerance of Helicobacter pylori, which enables it to colonize the gastric epithelium\textsuperscript{145}. Only a minor proportion of the microbiota has been cultured so far, thus we only have a fairly limited understanding of the bacterial diversity of the human body. However, modern techniques such as RT-PCR of 16S rRNA and proteomics are providing a greater understanding of the bacterial composition and diversity within the human body.

The oral cavity harbors up to 500 different bacterial species located on the teeth, gingival crevices, plaques, buccal mucosa and tongue\textsuperscript{144}. The main phyla found are Deferrribacteres, Spirochetes, Fusobacteria, Actinobacteria, Firmicutes, Proteobacteria, and Bacteroidetes\textsuperscript{144}. The
bacteria attached to the tooth surfaces form biofilms known as dental plaques. In the gingival crevices large amounts of Gram-negative anaerobes, especially *Porphyromonas gingivalis*, are believed to be involved in the pathogenesis of periodontal disease [149].

The environment of the gastric sac is highly acidic, and thus acts as a chemical tool to limit the local bacterial flora and entry of pathogens into the intestine. Still, some organisms are able to survive the acidic environment, and more than a hundred different species have been found here [11]. *H. pylori* is a known causative agent of gastric and duodenal ulcers and also of gastric cancers [145]. However, it is commonly found to colonize the gastric epithelia of healthy individuals as well [145].

A great variation of microbial colonization is seen in the various sections of the intestine. Few bacteria are found in the duodenum and jejunum, whereas the ileum contains up to $10^9$ bacteria / ml lumen content with a great degree of species variation. However, the richest and most diversified bacterial population is found in the human colon [197]. Most of the bacteria are strict anaerobes with the most abundant phyla being Firmicutes and Bacteroidetes, followed by Proteobacteria, Actinobacteria, Fusobacteria, and Verrucomicrobia [41].

The urogenital tract is kept mostly sterile by the flushing effect of urine. The main colonizers of the vaginal epithelium are *Lactobacillus*. In fact, in some individuals various
Lactobacillus species were found to be the only microbes present. However, other common colonizers were *Gardnerella vaginalis* and streptococcal species\[^{70}\].

Unlike the above-described mucosal surfaces, the respiratory tract is equipped with efficient mechanical tools, such as cilia-mediated movement of the mucus, keeping the trachea, bronchi and alveoli sterile under normal healthy conditions\[^{211}\]. However, the upper parts of the respiratory tract, such as the nose, nasopharynx and oropharynx are inhabited by a great variety of microbes. These include staphylococci, streptococci, Corynebacteria and Gram-negative cocci. Some of the colonizing bacteria, for example *Streptococcus pneumoniae* and *Neisseria meningitidis*, may cause life-threatening infections such as pneumonia and meningitis\[^{36,197}\].

**Immunology of the mucosal surface**

The immune system of the mucosal surfaces is different from the systemic immune system. Both harmful antigens, such as those of pathogens, and non-harmful antigens, such as degraded food and components of commensal bacteria, are present in the mucosa. An equal immune response to all types of antigens could be harmful to the human host. Thus induction and maintenance of tolerance to many bacteria is essential. The polarized epithelium operates together with the underlying APCs to monitor the microbial colonization. These cells carry receptors on their surfaces, including those for C components, antibodies and lipopolysaccharide (LPS).
A specific system of pattern recognition molecules (PRMs) such as the Toll-like and Nod-like receptors (TLRs and NLRs, respectively) can give immunosuppressive or immunoinductive signals depending on where and by which factors the receptors are engaged. In general, luminal antigens cause no harm, but antigens on the basolateral side of the epithelia may cause immunological activation[163]. The adaptive branch of the immune system has very special features at the mucosa, especially in the intestine. Both diffuse and well-structured lymphoid tissues, such as Peyer’s patches, exist in direct connection to the mucosal epithelium[22]. There is a predominance of memory lymphocytes in the tissue and a constant secretion of IgA – the most abundant immunoglobulin of the mucosal surfaces[23, 24].

A key difference in the immunological decision of tolerance or activation is the environment in which the APCs meet their antigens, and present them to the T-cells. In the absence of inflammatory stimulation, CD103-positive dendritic cells will induce a regulatory T-cell phenotype[163]. However, if the mucosal barrier is breached, e.g. by infecting microbes, a multitude of innate immune molecules are activated thus altering the nature of the antigen presentation towards a protective adaptive immune response[178]. During neonatal life, an adaptive immune response is not yet fully developed, and the function of the innate immune system at the mucosal surfaces is therefore particularly important for the health of the newborn[96].
When the barrier is breached – encounter with the complement system

Complement activation
The complement system is a complex enzymatic cascade that has evolved to both complement the immunological processes in the body but also to orchestrate the precise targeting of these processes. Complement is activated after binding to specific surfaces. Recognition molecules bind to exposed foreign or altered-self molecular structures, including bacteria, viruses, antibody-antigen complexes and apoptotic cells. In principle C components are capable of targeting every surface in the human body. However strict regulation of the activation ensures that C is only activated when needed – at least in healthy individuals, reviewed in [40, 117, 162].

The C cascade is divided into three different pathways, the classical, the alternative and the lectin pathway (Figure 2). The classical and lectin pathways are activated in very similar ways by the binding of specific soluble PRMs to their ligands [202]. The PRM of the classical pathway is C1q and those of the lectin pathway are the mannose binding lectin (MBL) and the ficolins H, L and M (also termed ficolins 1, 2 and 3). Binding of the PRMs to their respective targets induces conformational changes that affect the C1q-associated serine proteases, C1r/s, and MBL-associated serine proteases 1 and 2 (MASP1 and MASP2), respectively [7, 149]. The serine proteases activate C4. Activation cleaves C4 into C4a and C4b, revealing a hidden thioester site, which
covalently links C4b to the target surface in close proximity to the activating complex. C4b binds C2, which is subsequently cleaved by C1s or MASP. The two cleaved components join to form the C4b2a complex also known as the C3-convertase of the classical and lectin pathways \cite{29,170}. This is a key step in C activation. The C4b2a complex activates C3 leading to deposition of C3b on the target surface and release of the anaphylatoxin C3a into the surrounding microenvironment \cite{202}.

C3 is a very unstable molecule in solution and auto-hydrolyzes readily into C3(H2O). This marks the initiation of the alternative C pathway. C3(H2O) exposes new binding sites and binds factor B, which in the presence of factor D is cleaved to Bb. The C3(H2O)Bb complex functions as a soluble C3 convertase producing C3b which subsequently binds covalently to nearby surfaces. On the surface, the C3b again binds factor B forming the surface-bound C3bBb complex, also known as the C3 convertase of the alternative pathway \cite{10}. The C3b formed can again bind new factor B molecules and form even more C3-convertases. Thus the alternative pathway functions as a very efficient amplification loop and can enhance C activation created by auto-hydrolysis or by utilizing the C3b formed by the classical and lectin pathways, reviewed in \cite{59}. 
Literature review

A) Activation

B) Amplification

C) Terminal pathway

Target surface
Figure 2: The complement system.

A) Activation. Classical pathway: C1-complex (light blue) binds to IgG/IgM (blue) or CRP/pentraxins (red) on the target surfaces. Lectin pathway: MBL/Ficolins-MASP complex (green) binds carbohydrate structures on the target surface (xxx). For the lectin and classical pathways, activation involves the cleavage of C4 and C2, which generates the C3 convertase C4b2a on the target surface. Alternative pathway: In the fluid phase C3 auto-hydrolyses to C3(H2O). It binds factor B, which in the presence of factor D is activated to form the C3(H2O)Bb complex.

B) Amplification. Surface-bound C4b2a and fluid-phase C3(H2O)Bb cleave C3 to C3a and C3b. C3b is deposited on the nearby surface where it binds factor B. Factor D cleaves factor B to form a surface-bound C3bBb complex, which is further stabilized by properdin (P, blue triangle). This complex is a C3 convertase enzyme able to cleave new C3 molecules thus amplifying the signal.

C) Terminal pathway. Newly created C3b can attract C5 to cleavage by C4b2a or C3bBb. C5 is cleaved to C5a and C5b. C5b binds C6, C5b-6 binds C7 and C8 whereby the complex is inserted into the cell membrane. Finally C9 is recruited to form a pore that allows exchange of ions and small molecules through the double phospholipid membrane.

The formation of C3 convertases (either C4b2a or C3bBb) on target surfaces pave the way for the initiation of the terminal pathway, shared by all three activation cascades. The C3 convertases bind C3b and form the C5 convertase (C4b2aC3b or C3bBb3b). C5 is cleaved leading to generation of C5b and release of the anaphylatoxin C5a in the fluid phase [117, 162]. The deposited C5b can now bind C6 and C7, whereafter the C5b-7 complex can bind to a membrane and recruit C8. Together they form a complex that is inserted into double phospholipid cell membranes. Finally multiple molecules of C9 are inserted into a ring-like structure forming the C5b-9 complex, or the membrane attack
complex (MAC). The ring structure is essentially a pore in the cell membrane allowing free movement of water and other solutes. Calcium and sodium influx into the cells causes activation of many intracellular processes. After having the cell surface covered with MAC the osmotic gradient of the cell is destroyed and it ruptures and dies \[132\].

**Complement regulation**

The unstable nature of C3 allows it to constantly probe the surfaces of the immediate surroundings. When the soluble C3bBb convertase is active, C3 is cleaved in the fluid phase and deposited on a nearby surface. On foreign or modified host surfaces this signal can be quickly multiplied. However, on our own healthy cells, several strong complement regulators exist and more can be recruited. Basically the complement regulation can be divided into five main mechanisms; inhibition of C3 convertase formation, factor I cofactor activity, decay-accelerating activity for the C3 convertase, inhibition of lysis and finally cleavage of anaphylatoxins, reviewed in \[40\]. Some regulators are found on the cell membrane of the human cells such as CD46, CD55, CD59, Complement receptor (CR) type 1 and CR of the immunoglobulin superfamily CRIg \[219\]. Others are found in the fluid phase and recruited to the C targeted cell surfaces. These include factor H that controls the amplification loop and C1 inhibitor and C4b binding protein (C4BP), which inhibit both the classical and the lectin pathway \[219\]. Factor I is a fluid phase serine esterase enzyme recruited to C3b and C4b on self surfaces by interaction with the regulators CD46,
CR1, C4BP and factor H. The binding leads to the cleavage of C3b first to iC3b, and subsequently to C3c and C3dg. C4b is degraded to C4c and C4d. CR1, C4BP and factor H inhibit the formation of new C3 convertases. CD55, CR1, C4BP and factor H have decay accelerating activity. Formation of the MAC complex is inhibited by CD59 on cell membranes and by vitronectin and clusterin in the fluid phase. Finally, the signaling properties of the fluid-phase anaphylatoxins, C3a and C5a, are influenced by cleavage with specific carboxypeptidases. The regulation of C has been reviewed in [40].

**Complement receptors**

Both the soluble and the surface-bound complement activation products have several receptors on various immune and non-immune cells. CR1 binds C4b and C3b deposited on e.g. microbial surfaces. CR1 also acts as a cofactor for factor I [93]. The iC3b formed after C3b inactivation by factor I is a ligand for CR3 and CR4 on phagocytes [199]. This enhances phagocytosis and in the case of CR3 helps to induce cytokine production, leukocyte trafficking and even synapse modeling [199]. CR2 is found on B-cells, where it interacts with both iC3b and the further degradation products C3dg and C3d [206]. Stimulation of this receptor acts as a co-stimulatory signal for B-cell activation [206]. C3a and C5a are bound by specific receptors (C3aR and C5aR1 and C5aR2, respectively) expressed e.g. by endothelial cells and many types of leukocytes [207]. Finally, even the initial PRMs such as C1q and MBL have receptors of
their own. However, to this date the specific interactions necessary for binding of these molecules to their receptors are only poorly understood [139].

**Functions of complement**

It can be deduced from the above that the physiological functions of C span far wider then just insertion of lytic pores into the membranes of foreign pathogens. Alas, only few examples exist of this being an effective killing mechanism of microbes, e.g. Neisseriae [155]. Instead inflammatory signaling and opsonization for phagocytosis are key elements of the anti-microbial functions of C [203]. At the site of infection where C is first activated, several of its components are bound by receptors on APCs. Both foreign and host cells and other structures coated with C3b and C4b become targets for phagocytosis by the dendritic cells and macrophages through direct binding to receptors on the surfaces of the phagocytes [199].

Although C is mainly considered a part of the innate immune response, it is becoming evident that C also acts as a bridge to adaptive immunity and even the coagulation system [114]. Simultaneous activation of the C3aR/C5aR and TLRs of APCs strongly enhances the danger-signaling and subsequent production of pro-inflammatory cytokines [218]. The stimulation of APCs coordinated by complement affects the activation of T-cells. However, in addition to modulating the APC-T-cell interactions, C has also been suggested to interact with receptors on the T-cells directly. T-cells respond by
altered proliferation and differentiation to ligand binding to surface receptors such as the anaphylatoxin receptors C3aR and C5aR and also the complement regulators CD46, CD55, CD59 and CR1 \[39\]. In addition, direct binding of C1q or C1q-coated immune complexes also affects the function of T-cells \[28, 76\].

Complement was initially named based on the observation that it “complements” the function of antibodies in cell killing. The C receptors CR1 and CR2 are found on the surfaces of B-cells. Here they function as co-receptors, e.g. the CR2-CD19-CD81 co-receptor complex, and allow stronger stimulation of B-cell receptors when they bind an antigen coated with C \[206\]. In addition, CR1 and CR2 on follicular dendritic cells in the germinal centers are responsible for long-term retention of antigens and thus stimulate the B memory cell generation \[40, 47\].

In addition to the C functions described above, the released anaphylatoxins, C3a and especially C5a are powerful effector molecules. They stimulate the local endothelium to induce integrin expression allowing the recruitment of lymphocytes to the tissue and local activation of the coagulation system, thus helping to contain a possible infecting pathogen \[114\]. Finally, they exert chemoattractant effects by binding to their respective receptors on neutrophils, monocytes and macrophages attracting the cells to the site of infection or inflammation, and subsequently activating them \[207\].
Importantly, complement does not only function in orchestrating an anti-microbial defense. It is also involved in the resolution of inflammation and removal of apoptotic cells. Complement is primarily activated on the surface of apoptotic cells through targeting by C1q, which leads to C3b and C4b deposition and subsequent phagocytosis[^48]. A key feature of this process is the clear difference in action seen in the very strong anti-microbial response and the more gentle response utilized in clearance of endogenous material. An example is C-reactive protein (CRP), which binds to apoptotic cell surfaces. CRP mediates the binding of C1q to the surface, and thus initiates complement activation. However, simultaneously, CRP binds factor H and thus inhibits activation of C3 and the terminal pathway[^72]. Dying cells may shed the membrane regulatory molecules such as CD46, CD55 and CD59 allowing C attack against the cell surface[^42]. However, residual presence of regulators on the surface and in the immediate surrounding fluid phase help to keep the activation at a minimum[^162, 219]. Finally C has been shown to affect tissue repair by neovascularization and to stimulate the mobilization of hematopoetic stem/progenitor cells from the bone marrow to replenish the pool of immune cells in blood[^71, 137].

**Complement and mucosal surfaces**

Although C is present mainly in blood, it is also found in serous exudates on mucosal surfaces, such as in the oral cavity or the airways[^16, 147]. This is particularly seen under pathological conditions, for example in mechanical damage
or during infection, e.g. periodontal disease\textsuperscript{[33]}. When serous exudates enter the mucosal surfaces, innate immune proteins interact with mucosal surface proteins, and together these molecules create a specific interplay of defense mechanisms against invading microorganisms.

**Placenta**

**Placental anatomy and development**

After the human egg is fertilized it travels through the Fallopian tube to the site of implantation in the uterus. During this journey mitotic divisions of the original zygote divide the early embryo into an inner cell mass, the embryoblast, and an outer peripheral cell layer, the trophectoderm. In time, the trophectoderm develops into the placental and amniotic membrane tissues\textsuperscript{[156]}. During the actual implantation the trophectoderm interacts with the maternal decidua through expression of several mediators such as cytokines, prostaglandins, selectins, integrins and many more\textsuperscript{[6]}. As the invasion proceeds the cytotrophoblasts of the trophectoderm fuse to form the syncytiotrophoblast layer (Figure 3). This cell layer has direct contact with the maternal tissue, and later with the maternal blood. The syncytium forms trabeculae that protrude from the developing fetus. These are eventually filled with fetal vessels (Fv) and form the placental villous trees\textsuperscript{[156]}. Initially, the fetus is nourished from the uterine glands. Specialized endovascular cytotrophoblasts invade further into the decidua early on, and block the maternal vessels.
Figure 3: Anatomy of the human placenta.

During placentation extravillous trophoblasts (green) invade into the maternal decidua. The outer cell layer fuses to form the syncytiotrophoblast layer (blue) and forms trabeculae, which anchor to the maternal tissue. In time these are filled with fetal vessels (Fv) and make up the structural unit of the placenta, the villous tree. After approximately 10-12 weeks of gestation the extravillous trophoblasts transform the maternal spiral arteries (SA) and open up the blood flow into the intervillous space (IVS). From here on the exchange of nutrients and waste products occurs across the syncytiotrophoblast layer. In case of syncytial damage, fibrinoids form at the barrier between fetal and maternal tissue (marked as XXX). Fibrinoids are also found at other locations, e.g. surrounding invading trophoblasts, at the anchoring sites of the villi and covering the chorionic plate.
Only after 10-12 weeks of gestation is the maternal blood flow released into the intervillous space of the placenta \cite{69}. As pregnancy continues, the villous trees evolve continuously. The syncytiotrophoblast layer is replenished from the proliferating cytotrophoblasts, while the aging nuclei of the syncytium are shed from the villous surface and into the maternal circulation. During pregnancy high amounts of fragmented syncytiotrophoblasts and cytotrophoblasts are found in the maternal circulation \cite{88}. Often this shedding along with other injuries of the syncytium leaves the fetal tissue exposed. These areas become covered in a fibrin-rich matrix called fibrinoid which is, amongst other things, used as a scaffold for re-epithelialization of the syncytium-void areas \cite{136}.

**Immunology of the placenta**

Because of the fetal expression of paternal antigens, the direct contact between fetal and maternal tissue poses a great immunological challenge. Thus, the maternal decidua is populated with a great number of immunological cells all displaying an immunosuppressive phenotype \cite{26}. Furthermore, the fetal tissue is enclosed in a coat of multinucleated trophoblast cells, the syncytiotrophoblast, which functions as a physical barrier, blocking migration of maternal leukocytes into the fetal tissue. Together with the invading extravillous cytotrophoblast cells these are the only fetal tissues in direct contact with the maternal immune system. Unlike most other eukaryotic cells the trophoblast
cells are negative for the major histocompatibility complex (MHC) class I (HLA-A, HLA-B) and class II (HLA-DP, -DQ, -DR) antigens, which, when carrying foreign peptides, are the targets for cytotoxic and T helper lymphocytes [68]. Instead, the trophoblast cells express HLA-G and HLA-E antigens that suppress activities of e.g. natural killer (NK) cells. In addition, the trophoblast cells actively secrete various immunosuppressive molecules such as progesterone, prostaglandins and cytokines IL-4 and IL-10 [68]. A disadvantage of the decreased immunological activity at the fetomaternal interface is an increased susceptibility to uterine and other maternal infections. This may progress to intra-amniotic infection often leading to preterm birth, and possibly other pregnancy complications [52].

**Complement in the placenta**

The placental syncytiotrophoblast, and in cases where the syncytiotrophoblast is damaged, the intravillous tissue is in a direct contact with maternal blood. Thus, the potential activation of the C system against the “foreign” fetal tissue poses a great threat to the fetus [108]. The syncytiotrophoblast layer of the placenta has a high turnover. As described above, C activation is involved in clearance of apoptotic cells. Therefore, it is not surprising that C activation is observed in a normal healthy pregnancy [161, 205]. This makes C regulation essential in the placenta [191]. Indeed, dysregulated C activation could contribute to placental damage in many pregnancy complications, including pre-eclampsia (PE). Early C activation was found
to be a predictor of PE [107]. Of specific interest to our study, increased deposition of the C5b-9 terminal complex on trophoblasts has been found associated with fibrinoid deposits at sites of villous injury, especially in complicated pregnancies such as PE and intra-uterine growth restriction (IUGR) [157].

**SALSA protein**

**Discovery**

The protein salivary scavenger and agglutinin (SALSA) was first described as a 300-400 kDa streptococcal agglutinating agent from saliva, and named salivary agglutinin (SAG) [45]. Later, others isolated a similar 340 kDa glycoprotein from bronchoalveolar lavage fluid after co-purification with surfactant protein D (SpD). This glycoprotein was named gp340 [65]. At the same time, another study found the gene for SALSA deleted in several brain tumors, suggesting a tumor-suppressor function for the protein, and named it ‘deleted in malignant brain tumors 1’ (DMBT1). The gene was accordingly named DMBT1 [119]. Only later was it realized that SAG and gp340 are variant proteins encoded by the same gene, DMBT1, in different tissues [66, 102, 152]. Because of the discrepancies of names mentioned above we suggested the name SALSA (Salivary scavenger and agglutinin) for this protein based on its initial discovery in saliva and its function as a scavenger and agglutinin.
**DMBT1 gene**

The *DMBT1* gene spans more than 80 kb of the human genome and contains at least 54 expressed exons \(^{[120]}\). The gene encodes 13 highly conserved scavenger receptor cysteine-rich (SRCR) domains. These 109 amino acid motifs are found as pearls on a string separated by SRCR interspersed domains (SIDs) (Figure 4). The stretch of 13 SRCR domains is followed by two C1r/C1s, urchin embryonic growth factor and bone morphogenetic protein-1 (CUB) domains encompassing a 14\textsuperscript{th} SRCR domain. Finally a zona pellucida (ZP) domain is found in the most C-terminal end \[^{[66, 120]}\]. However, the actual expressed *DMBT1* mRNA and protein products have been found to display a great variation from the full gene. The *DBMT1* gene is built up of specific locus repeats of 3-4 kb with up to 99.8 % homology. These repeats each encode a SRCR domain and a SID domain. The highly repetitive sequence is believed to facilitate alternative splicing leading to expression of several SALSA protein isoforms \(^{[125]}\). This is indeed observed in several cases with mRNA transcripts encoding between 8-13 of the N-terminal SRCR domains \[^{[66, 120, 124]}\]. Variations in the amounts of expressed repeats have been found in up to 28 % of normal individuals \(^{[121]}\). All in all the genetic analyses so far reveal at least seven distinct alleles. These are present in a highly unstable genetic region, where the loss of heterogeneity contributes to making these variations more readily expressed \[^{[125, 213]}\].
Figure 4: SALSA structure.

Shown is the domain structure of SRCR superfamily (SRCR-SF) proteins, including SALSA and SALSA orthologs. Some proteins are found in secretions while others contain a transmembrane region and are anchored into the cell membrane.

Glycosylation of the SALSA protein

It has been estimated that up to 25-45 % (w/w) of SALSA is carbohydrate \cite{45, 66}. SALSA contains all the major sugar components including glucose, galactose, fucose, mannose, N-acetylglucosamine, N-acetylgalactosamine, sialic acid and many of their derivatives \cite{46, 61, 141}. Sequence analysis
revealed 14 putative N-glycosylation sites and a number of O-glycosylation sites situated primarily in the SIDs [66, 180]. Based on this it has been suggested that the heavy glycosylation may force the SIDs into an extended conformation as observed for mucins [166]. Thus alternating stretches of SIDs and globular SRCR domains would constitute the structural layout of SALSA. In addition to the above-described variations in SALSA mRNA transcripts from the DMBT1 gene, further protein variation is seen with the glycosylation patterns. Differences in saliva-derived SALSA were found to correlate to secretor (Se(+)) and non-secretor (Se(-)) status (± expression of the α1-2fucosyl-transferase). The blood group antigens, ABO, and Lewis antigens b and y (Leb and Ley) were found on SALSA from Se(+) individuals. In contrast, SALSA from Se(-) individuals did not contain ABO, Leb nor Ley antigens. Instead Lewis antigens a and x (Lea and Lex) were present [46, 98]. In addition, the SALSA glycosylation patterns varied further in different secretions [46, 180]. Finally, variations correlating to hormonal changes during the menstrual cycle were found in the MECA-79 carbohydrate epitope, which is also found on SALSA [153]. Taken together, SALSA is a highly glycosylated protein. The specific glycosylation patterns vary between individuals and even with hormonal variations. Furthermore, SALSA expressed in different tissues also displays varying glycosylation.
SALSA across species
The SALSA protein consists of SRCR, CUB and ZP domains. All of these are evolutionarily highly conserved units and present in a multitude of proteins with varying functions [18, 19, 79, 80, 176]. SALSA orthologs have been described in the mouse (CRP-ductin, vomeroglandin, muclin, apactin), rat (ebnerin), rabbit (hensin), cow (bovine gall bladder mucin), pig (porcine dmbt1) and rhesus monkey (H3) [1, 30, 35, 56, 97, 115, 138, 187, 189]. The described proteins show variation in the number of SRCR and CUB domains, but all contain only one C-terminal ZP domain (Figure 4) [100, 101].

SALSA orthologs have been found expressed at various epithelial surfaces such as in the mouse pancreas and intestine, as well as in monkey and rat uterine epithelium [1, 30, 35, 110, 198]. The SALSA orthologs were involved in bacterial binding, actin skeleton remodelling and pheromone and taste perception [97, 110, 115, 189]. The rabbit ortholog for SALSA, hensin, was shown to directly influence the differentiation of stem cells and epithelial cells [188]. The specific mechanism of hensin function involves hensin polymerization through interactions with galectin-3 and integrins followed by incorporation into the extracellular matrix (ECM) [63, 64, 200].

SALSA expression in humans
The animal data above indicate that SALSA is expressed in a variety of tissues throughout the body. The same holds true for SALSA expression in humans. RT-PCR analyses originally showed that the main tissues of expression are the lung,
Literature review

trachea, salivary glands, small intestine and stomach \[66,119\]. These are all tissues with a mucosal surface. At the protein level further analysis by immunohistochemistry identified SALSA in secretory glands and associated with the epithelial layer in the lungs, oral cavity, trachea, gastrointestinal tract, pancreas, mammary glands, vagina and cervix \[21, 66, 83, 121, 122, 185, 195\]. In addition to the mucosal surfaces and secretory tissues, SALSA has been found in the heart, liver, skin and tonsils \[122, 129\]. Several soluble forms of SALSA have been found in body fluids lining the mucosal surfaces such as saliva, tear fluid, respiratory mucosal secretions and pancreatic juice \[45, 54, 65, 180, 195\]. Notably, SALSA has not been observed in human blood or plasma. For an overview of the current knowledge of SALSA expression, see Table 1.

To date, limited quantitative data on the levels of secreted SALSA is available. One of the early studies estimated that SALSA makes up less than 0.5 % of the total protein content of saliva. As the salivary protein amount has been estimated to be 1.4 ± 0.2 mg/ml, this would mean approximately 7 μg /ml for SALSA concentration \[45, 183\]. This was supported by later studies suggesting a concentration of 1-10 μg/ml \[185\]. Analysis of SALSA in tear fluid describes it to be more abundant than mucins, but without more specific data \[180\].

SALSA is highly expressed in proliferating and developing epithelia \[85, 121, 123\]. In addition, an induced expression of SALSA in infected and inflamed tissues has been observed. In the gut increased expression was found to associate with inflammatory bowel disease and especially Crohn's disease.
Immunohistochemistry showed only sporadic staining of SALSA in the connective tissue, endothelial cells and myocytes in healthy myocardium. However, in cases of endocarditis, SALSA expression was greatly increased in the local areas of bacterial colonization. Specifically, granulocyte-depleted fibrin formations were positive. Finally, SALSA was found weakly expressed by biliary epithelial cells in the liver, but increased expression was observed under certain pathological conditions such as hepatolithiasis.

SALSA was early on described as a tumor suppressor in malignant brain tumors. Since then altered expression has been found in a number of tumors. Genetic variations were found in up to 89% of tested lung and brain cancer cell lines. Healthy tumor-flanking tissue had elevated levels of SALSA compared to healthy tissue. However, down-regulation of SALSA was observed in lung and mammary gland tumors compared to the tumor-flanking tissue. Similarly, salivary gland tumors were found to have decreased SALSA expression compared to both normal tumor-flanking tissue and healthy tissue. In the esophagus, squamous cell carcinomas had reduced SALSA expression, whereas adenocarcinomas appeared to have an increased expression of SALSA.
Table 1: SALSA expression in human tissues and secretions

<table>
<thead>
<tr>
<th>Tissue</th>
<th>Location</th>
<th>Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Adult</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone marrow</td>
<td>Not known</td>
<td>RT-PCR [121]</td>
</tr>
<tr>
<td>Brain</td>
<td>Purkinje cells, granular layer, some astrocytes and neuronal cells, perivascular macrophages</td>
<td>RT-PCR, IHC [66, 119, 121]</td>
</tr>
<tr>
<td>Colon</td>
<td>Epithelium</td>
<td>IHC [122]</td>
</tr>
<tr>
<td>Duodenum</td>
<td>Enterocytes, crypt cells, Brunner’s glands</td>
<td>IHC [122]</td>
</tr>
<tr>
<td>Esophagus</td>
<td>Epithelium, mucus glands</td>
<td>IHC [122, 123]</td>
</tr>
<tr>
<td>Eye</td>
<td>Acinar cells of lacrimal glands, conjunctival epithelium</td>
<td>RT-PCR, IHC [81]</td>
</tr>
<tr>
<td>Female genital tract</td>
<td>Epithelium</td>
<td>IHC [185]</td>
</tr>
<tr>
<td>Gallbladder</td>
<td>Epithelium</td>
<td>IHC [123]</td>
</tr>
<tr>
<td>Heart tissue</td>
<td>Endothelium, myocytes</td>
<td>RT-PCR [129]</td>
</tr>
<tr>
<td>Kidney</td>
<td>Collecting duct epithelium</td>
<td>RT-PCR [66]</td>
</tr>
<tr>
<td>Liver</td>
<td>Hepatocytes</td>
<td>IHC [122]</td>
</tr>
<tr>
<td>Lung</td>
<td>Epithelium, macrophages</td>
<td>RT-PCR, IHC [66, 119]</td>
</tr>
<tr>
<td>Lymph node</td>
<td>Not known</td>
<td>RT-PCR [121]</td>
</tr>
<tr>
<td>Mammary gland</td>
<td>Epithelium</td>
<td>RT-PCR, IHC [21, 66]</td>
</tr>
<tr>
<td>Pancreas</td>
<td>Langerhans islets, exocrine acinar cells, pancreatic ducts</td>
<td>RT-PCR, IHC [66, 122, 122]</td>
</tr>
<tr>
<td>Prostate</td>
<td>Not known</td>
<td>RT-PCR [66]</td>
</tr>
<tr>
<td>Tissue Type</td>
<td>Epithelial Type</td>
<td>Method(s)</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Salivary glands</td>
<td>Serous glands</td>
<td>RT-PCR, IHC</td>
</tr>
<tr>
<td>Sinus</td>
<td>Epithelium</td>
<td>RT-PCR, IHC</td>
</tr>
<tr>
<td>Skin</td>
<td>Epithelium</td>
<td>IHC</td>
</tr>
<tr>
<td>Small intestine</td>
<td>Epithelium, gastric glands</td>
<td>RT-PCR, NB, IHC</td>
</tr>
<tr>
<td>Spleen</td>
<td>Not known</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>Stomach</td>
<td>Epithelium, serous glands</td>
<td>RT-PCR, IHC</td>
</tr>
<tr>
<td>Testis</td>
<td>Not known</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>Thymus</td>
<td>Not known</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>Tonsils</td>
<td>Epithelium</td>
<td>IHC</td>
</tr>
<tr>
<td>Trachea</td>
<td>Serous cells</td>
<td>RT-PCR, IHC</td>
</tr>
<tr>
<td>Uterus</td>
<td>Not known</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>Fetus</td>
<td>Epithelium</td>
<td>IHC</td>
</tr>
<tr>
<td>Kidney</td>
<td>Collecting duct epithelium</td>
<td>IHC</td>
</tr>
<tr>
<td>Lung</td>
<td>Not known</td>
<td>NB</td>
</tr>
<tr>
<td>Skin</td>
<td>Epithelium, epidermis</td>
<td>IHC</td>
</tr>
<tr>
<td>Intestine</td>
<td>Epithelium</td>
<td>IHC</td>
</tr>
<tr>
<td>Kidney</td>
<td>Collecting duct epithelum</td>
<td>IHC</td>
</tr>
<tr>
<td>Lung</td>
<td>Not known</td>
<td>NB</td>
</tr>
<tr>
<td>Skin</td>
<td>Epithelium, epidermis</td>
<td>IHC</td>
</tr>
<tr>
<td>Secreions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bronchoalveolar lavage</td>
<td></td>
<td>WB</td>
</tr>
<tr>
<td>Pancreatic juice</td>
<td></td>
<td>MS</td>
</tr>
<tr>
<td>Respiratory mucosal secretions</td>
<td></td>
<td>WB</td>
</tr>
<tr>
<td>Saliva</td>
<td></td>
<td>MS</td>
</tr>
<tr>
<td>Sinus mucosa</td>
<td></td>
<td>WB</td>
</tr>
<tr>
<td>Tear fluid</td>
<td></td>
<td>MS</td>
</tr>
</tbody>
</table>

IHC, Immunohistochemistry; NB, Northern blotting; WB, Western blotting; RT-PCR, Reverse transcriptase polymerase chain reaction; MS, Mass spectrometry.
Based on the location and differential expression described above, it has been suggested that SALSA functions in epithelial homeostasis, immunological surveillance and tumor suppression \cite{84, 100, 111}.

**Functions of SALSA**

**SALSA and epithelial differentiation**

*SALSA expression in non-differentiated cells*

By comparing SALSA levels between adult and fetal tissues, e.g. skin and the intestine, both an induced expression and a different subcellular localization were found in the fetal tissues \cite{121, 123}. Together with its location in the proliferating intestinal and gastric mucosa, this finding suggests a role for SALSA in the development of epithelial tissues \cite{85, 121}. SALSA was primarily found in stem/progenitor cells rather than in terminally differentiated cells \cite{30, 83, 121}. Analysis of the gastric adenocarcinoma cell line AGS revealed that phorbol 12-myristate 13-acetate (PMA) stimulation led to down-regulation of SALSA expression. It was shown that SALSA expression correlated with the change from a proliferating phenotype towards a cell-differentiating phenotype. Subsequently, SALSA expression was down-regulated again prior to terminal differentiation \cite{83, 85}. The involvement in the regulation of the epithelial cell-cycle is further substantiated by observations of altered regulation of SALSA in a number of epithelial cancers \cite{21, 123}. 
SALSA and cell polarization
In Japanese, hensin, the name of the SALSA ortholog in rabbit, means “change in shape”. Extensive work has been performed to elucidate the mechanisms involved in rabbit kidney epithelial cell differentiation, and hensin was found to play a key role \[^{[3]}\]. It was shown that hensin is secreted as a monomer. It then polymerizes and deposits into the ECM \[^{[63]}\]. The polymerization occurs through interactions with at least three other proteins: integrins, cyclophilin A and galectin-3 \[^{[64, 146, 201]}\]. The suggested mode of action was that hensin through its SRCR domains binds to integrins on the cell surface. Galectin-3 and cyclophorin A then mediate the polymerization and deposition into the ECM of hensin, which finally allows for the signalling to the cell. This provides an insight into how SALSA in humans may function in the epithelial differentiation.

Functions of SALSA in the innate immune system

Interactions with endogenous molecules
It has been observed that native SALSA found in saliva and pulmonary secretions polymerizes into large complexes \[^{[111, 141, 216]}\]. SALSA has been found to interact with numerous endogenous proteins such as surfactant proteins A (SpA) and D, secretory IgA, lactoferrin, fibrin/fibrinogen, trefoil factor 2, mucin-5B and C1q \[^{[17, 65, 129, 142, 172, 194-196]}\]. A list of SALSA ligands is shown in Table 2. The native form of SALSA in saliva is found in a complex with IgA \[^{[45]}\]. The interaction between IgA and SALSA was calcium-dependent and was shown to mediate the agglutination of Streptococcus mutans \[^{[99, 141, 172]}\].
Table 2: Endogenous ligands of SALSA.

<table>
<thead>
<tr>
<th>Interaction partner</th>
<th>Interaction site</th>
<th>Functional relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>SpD</td>
<td>SpD globular domain</td>
<td>Microbial agglutination [65]</td>
</tr>
<tr>
<td>SpA</td>
<td>SpA globular domain</td>
<td>Microbial agglutination [196]</td>
</tr>
<tr>
<td>IgA</td>
<td>SRCRP2 of SALSA</td>
<td>Microbial agglutination [45]</td>
</tr>
<tr>
<td>C1q</td>
<td>Globular domain of C1q</td>
<td>C activation [17]</td>
</tr>
<tr>
<td>Lactoferrin</td>
<td>SRCRP2 of SALSA</td>
<td>Mucosal surfaces, bacterial binding [118]</td>
</tr>
<tr>
<td>DNA</td>
<td>Phosphate groups of DNA</td>
<td>Inflammation [44]</td>
</tr>
<tr>
<td>Heparan sulfate</td>
<td>Sulfated groups of heparan sulfate</td>
<td>Inflammation [44]</td>
</tr>
<tr>
<td>Trefoil factors</td>
<td>Not known</td>
<td>Tissue homeostasis [194]</td>
</tr>
<tr>
<td>MUC5B</td>
<td>Not known</td>
<td>Microbial agglutination [210]</td>
</tr>
<tr>
<td>Fibrin</td>
<td>Not known</td>
<td>Not known [129]</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>Not known</td>
<td>Not known [129]</td>
</tr>
<tr>
<td>Erythrocytes</td>
<td>Not known</td>
<td>Aggregation [129]</td>
</tr>
<tr>
<td>Platelets</td>
<td>Not known</td>
<td>Aggregation [129]</td>
</tr>
</tbody>
</table>

SRCRP2, the bacterial binding peptide sequence of the SRCR domains, see below.

The IgA binding site on SALSA was identified as the bacterial binding SRCR peptide 2 (SRCRP2) although IgA and S. mutans did not appear to be competing for SALSA binding [99]. The complex formation between SALSA and IgA has an
enhancing effect on bacterial agglutination, but the main binding ability was ascribed to SALSA\textsuperscript{[172]}.

The C1q molecule consists of 18 polypeptide chains held together by disulfide bridges to form a bouquet-like structure. Each chain contains a collagen domain (stalk) and a carboxy-terminal globular domain (the flowers)\textsuperscript{[159, 181]}. The globular domain is used for target recognition while the collagen domains bind the enzymes C1r and C1s. Conformational changes occurring by ligand binding of the globular regions lead to activation of C1r/s followed by activation of the classical pathway of C\textsuperscript{[49, 202]}.

C1q binds specifically the immunoglobulin classes M and G, however, other endogenous non-immunoglobulin ligands have been found, e.g. SALSA\textsuperscript{[17]}. C1q is a key PRM known to interact with microbial ligands directly, such as LPS of \textit{Salmonella} Typhimurium, but also through interaction with immunoglobulins bound on the pathogen surface\textsuperscript{[162, 171]}. In addition, an important function of C1q is clearance of apoptotic cells\textsuperscript{[190]}. Detailed studies have shown that C1q binds SALSA through the globular domain in a region close to the immunoglobulin binding site.\textsuperscript{[91]} It was suggested that the interaction between C1q and SALSA could activate the C system\textsuperscript{[17]}.

SALSA also interacts directly with the two collagen-containing lectins SpA and SpD\textsuperscript{[65, 196]}. The interactions are mediated through the carbohydrate binding globular domain of SpD and SpA and as for IgA, these interactions are calcium-dependent\textsuperscript{[65, 196]}. As described below, SALSA
agglutinates bacteria. SpD and SpA are also molecules expressed at the mucosal surfaces with known bacterial agglutinating functions [57]. It has been shown that the three proteins have a co-operative effect against influenza A virus (IAV) [208]. This was shown by viral agglutination, hemagglutinin inhibition and viral neutralization assays [208].

Lactoferrin is an iron-chelating protein sharing many properties with SALSA. Lactoferrin is expressed on mucosal surfaces, and has been suggested to be involved in cellular growth and differentiation, in defense against bacterial and viral infection, and in dampening inflammation [204]. SALSA has been shown to interact directly with both the human and bovine form of the protein [43,118]. Lactoferrin mediates its immune functions directly by chelating iron from the local mucosal environment, thus depriving the microorganisms of this essential factor. Furthermore, it has been shown that proteolytic degradation of lactoferrin leads to the release of antimicrobial peptides [53,182].

Another important defense molecule at the mucosal surfaces is the mucin MUC5B. This protein forms large protein aggregates in the mucosal secretions and inhibits bacterial colonization. SALSA has been shown to bind MUC5B in both respiratory secretions and saliva [195,210]. It is currently not known what the in vivo outcome is of the interplay between these two molecules at the mucosal surfaces. SALSA has so far not been described free in serum. However, in addition to the interaction with plasma C1q, it has been described that SALSA binds to both fibrin and fibrinogen.
Furthermore, SALSA was found to bind to platelet and erythrocyte surfaces, and directly mediate erythrocyte aggregation \cite{129}. No further studies have been done on SALSA in cardiovascular diseases. However, the endogenous ligand SpD has been suggested to influence the development of atherosclerosis in mice \cite{104}.

**SALSA in infection and inflammation**
The very first isolation of SALSA from saliva was done using *S. mutans* \cite{45}. Since then a great number of bacterial species has been added to the list of SALSA binding partners. Based on the broad bacterial binding properties of SALSA, it has been described as a PRM, however, no actual receptor functions have been described yet \cite{44,100,104}. SALSA binds both Gram-positive and Gram-negative bacteria, including *Bifidobacterium lactis, Staphylococcus aureus, H. pylori, Actinomyces odontolyticus, Salmonella enterica* serovar Typhimurium, several lactobacillus strains and as well a long list of streptococcal species (Table 3) \cite{62,81,103,152,168}. Furthermore, the viruses IAV and human immunodeficiency virus type 1 (HIV-1) have been shown to bind SALSA \cite{60,133}. In addition, orthologs of SALSA in other mammalian species have been shown to bind *Haemophilus influenzae, Klebsiella oxytoca* and *Streptococcus pneumoniae* \cite{110}. The broad bacterial binding property of SALSA has been assigned to one particular peptide sequence within each SRCR domain (RVEVLYxxxSW), designated SRCR-peptide 2 (SRCRP2) \cite{13,94}. The presence of several SRCR domains in SALSA possibly allows for a multivalent binding interaction with bacteria.
Table 3: Detected microbial ligands of human SALSA

<table>
<thead>
<tr>
<th>Microbe</th>
<th>Specific strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lactobacillar strains:</td>
<td><em>L. rhamnosus, L. casei, L. reuteri, L. lactis</em>[^62]</td>
</tr>
<tr>
<td>Other bacteria:</td>
<td><em>Staphylococcus aureus, Bifidobacterium, Actinomyces, Salmonella enterica serovar</em></td>
</tr>
<tr>
<td></td>
<td>Typhimurium, <em>Helicobacter pylori, Haemophilus influenzae, Klebsiella oxytoca</em>[^39, 62, 81, 103, 110, 152, 168]</td>
</tr>
<tr>
<td>Viruses:</td>
<td>HIV, IAV[^60, 133]</td>
</tr>
</tbody>
</table>

The microbial ligands found to bind either human SALSA, the murine ortholog of SALSA or the recombinantly expressed peptide SRCRP2.

The above mentioned binding studies suggest that SALSA simply functions by agglutinating microorganisms at the mucosal surfaces, thus keeping them from infecting the tissue, as has been observed for *S. enterica*, HIV-I and IAV[^60, 133, 168, 214]. However, the role of SALSA is more complicated than that. In the body SALSA is found both adhered to e.g. epithelial and tooth surfaces and secreted into the fluid phase[^103]. The epithelium-attached localization of a protein with a solely bacterial agglutinating function would not appear to be beneficial for the human host. This paradox has been made clear by studies showing that SALSA in some
cases may be utilized directly by invading microbes. A study of dental caries showed that certain SALSA-phenotypes correlated positively with *S. mutans* adhesion to SALSA-coated hydroxyapatite surfaces and the development of dental caries. Other SALSA phenotypes displayed the opposite correlation \(^7^8\). In the case of HIV-1 infection, the salivary secreted SALSA protein was found to interfere with oral transmission. However, SALSA expressed on the vaginal epithelia had an enhancing effect on the infectivity of the virus \(^1^8^5\). These findings suggest that some microbes have evolved mechanisms to utilize SALSA to infect the human body.

Interestingly, the specific bacterial binding ability of SALSA has been found to depend not only on the phenotype of the protein but also on the location of the protein. Fluid-phase SALSA can bind and aggregate some streptococcal strains, while SALSA coated to a hydroxyapatite surface does not. Other species were only bound by surface-associated SALSA. This variation was observed for strains even within the same streptococcal species \(^1^0^3\).

In addition to the specific localization of SALSA, it seems that also the temporal expression may be crucial for the immunological function of the protein. Studies have shown that SALSA can be strongly induced by various immunological stimuli. In several chronic inflammatory conditions, the expression of SALSA was found to be upregulated \(^8^9, ^1^6^0\). The increased levels of SALSA in the intestinal epithelium of patients with inflammatory bowel
Literature review

disease, including Crohn’s disease, and in the ethmoid sinus mucosa of patients with chronic sinusitis suggest that SALSA expression is part of the mucosal inflammatory response [58, 89, 160]. In addition, a study of preterm infants revealed a clear increase in the pulmonary SALSA levels as a response to neonatal infection [128].

Further studies have been done to investigate the involvement of SALSA in the mucosal epithelial inflammatory response. PMA and dextran sulfate sodium (DSS) are chemicals used to incite epithelial cell damage to mimic pathologic inflammatory conditions. PMA stimulation of the pulmonary epithelial cell line A549 led to increased expression of SALSA. This coincided with induction of the inflammatory cytokines IL-6 and IL-8 [83]. DSS was used to stimulate colitis in a mouse model. Here increased SALSA levels were found specifically in the damaged intestinal epithelium [160]. Similar enhanced expression of SALSA was observed by epithelial cells after stimulation with the proinflammatory stimuli LPS, muramyl dipeptide and tumor necrosis factor α (TNF-α) [168]. It was further shown that the signalling cascade utilized nucleotide-binding oligomerization domain 2 (NOD2) and TLR4 activation of NFκB to alter DMBT1 gene activity. Interestingly, a feedback loop seemed to exist, as secreted SALSA was able to inhibit the epithelial response to S. enterica-derived LPS and muramyl dipeptide [168].
Aims of the study

The overall aim of this thesis work was to investigate the function of the SALSA protein. Based on the expression of the protein we sought to understand how SALSA interacts with the innate immune defense system and exerts its role at the mucosal surfaces and in tissues.

The specific aims were:

1. To analyze the expression of SALSA in tissues and secretions related to early life (II, III)
2. To identify novel ligands of SALSA (I, III)
3. To investigate interactions between SALSA and the complement system (I, III)
4. To understand the functional relevance of SALSA in the placenta and amniotic fluid (II, III)
Materials and Methods

The methods used in this work are described in detail in the original publications I-III. However, in the following section a general overview of the techniques will be given. The used techniques, biological materials, proteins, antibodies, and buffers are listed in Tables 4-7.

Table 4: Overview of techniques used in this thesis

<table>
<thead>
<tr>
<th>Technique used</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDS-PAGE and Western blotting</td>
<td>I, II</td>
</tr>
<tr>
<td>ELISA</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Flow cytometry</td>
<td>I</td>
</tr>
<tr>
<td>Complement activation assays</td>
<td>I</td>
</tr>
<tr>
<td>Glycoprotein and lectin staining</td>
<td>II</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>II</td>
</tr>
<tr>
<td>SALSA-bacterial binding assay</td>
<td>II</td>
</tr>
<tr>
<td>Immunohistochemistry</td>
<td>III</td>
</tr>
<tr>
<td>Coagulation assays</td>
<td>III</td>
</tr>
</tbody>
</table>
Materials and methods

Human samples

Amniotic fluid

Amniotic fluid (AF) samples were collected at the Obstetrics and Neonatology Units of Hospital Universitario Doce de Octubre, Madrid, Spain (n = 9), and at the Women’s Clinic of the Helsinki University Hospital, Helsinki, Finland (n = 98). Term AF samples were collected in the third trimester by amniocentesis or needle aspiration during caesarean section (n = 46) or vaginal (n = 27) delivery. In some cases mode of delivery was not registered (n = 34). The samples were collected from healthy controls and women diagnosed with PE, IUGR, diabetes mellitus type 1 (DM) and gestational diabetes mellitus (GDM). In addition, samples were collected in second trimester (17.2 ± 2.5 gestational weeks). These are referred to as early pregnancy samples. After collection, samples were immediately frozen and stored at -20°C.

Intestinal samples

Meconium (n = 9) and fecal (n = 9) samples were collected from healthy term newborns at the Obstetrics and Neonatology Units of Hospital Universitario Doce de Octubre, Madrid, Spain. Meconium was collected within the first 2 hours from birth and before feeding was started. Fecal samples were collected one week after birth. Prior to comparison of AF, meconium and fecal samples, proteins were extracted as described [92]. Dried AF or thawed meconium and fecal samples were re-suspended in PBS and
### Table 5: Human fluids and tissues used in this thesis

<table>
<thead>
<tr>
<th>Material</th>
<th>Origin</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal human serum (NHS)</td>
<td>Pool from healthy volunteers</td>
<td>I</td>
</tr>
<tr>
<td>MBL-deficient serum</td>
<td>Obtained from an individual found to lack MBL in a screen of healthy volunteers</td>
<td>I</td>
</tr>
<tr>
<td>MgEGTA-serum</td>
<td>NHS containing 10 mM EGTA and 5 mM MgCl₂</td>
<td>I</td>
</tr>
<tr>
<td>Heat-inactivated serum (HIS)</td>
<td>NHS incubated at 56 °C for 30 min.</td>
<td>I</td>
</tr>
<tr>
<td>Citrated plasma</td>
<td>Plasma collected in citrate-containing tubes and pooled (n ≥ 2)</td>
<td>III</td>
</tr>
<tr>
<td>Saliva</td>
<td>Pool from healthy volunteers</td>
<td>I, II</td>
</tr>
<tr>
<td>Amniotic fluid</td>
<td>Collected at the Obstetrics or Neonatology Units of Hospital Universitario Doce de Octubre, Madrid, Spain.</td>
<td>II</td>
</tr>
<tr>
<td>Amniotic fluid</td>
<td>Collected at the Women’s Clinic of Helsinki University Hospital, Helsinki, Finland</td>
<td>III</td>
</tr>
<tr>
<td>Meconium</td>
<td>Collected at the Obstetrics or Neonatology Units of Hospital Universitario Doce de Octubre, Madrid, Spain.</td>
<td>II</td>
</tr>
<tr>
<td>Feces</td>
<td>Collected at the Obstetrics or Neonatology Units of Hospital Universitario Doce de Octubre, Madrid, Spain.</td>
<td>II</td>
</tr>
<tr>
<td>Placenta (frozen tissue)</td>
<td>Obtained from the Finnish Genetics of Preeclampsia Consortium (FINNPEC)</td>
<td>III</td>
</tr>
<tr>
<td>Placenta (paraffin embedded tissue)</td>
<td>Obtained from Medical University of Graz, Austria</td>
<td>III</td>
</tr>
</tbody>
</table>
Materials and methods

The samples were then subjected to a FastPrep 24 (MP Biomedicals, California, USA) according to the manufacturer's instructions. The resulting protein extracts were stored at -80°C.

Placental samples
Paraffin embedded placental samples were obtained from the Department of Obstetrics and Gynaecology, Medical University Graz, Austria and frozen sections from the Finnish Genetics of Preeclampsia Consortium (FINNPEC) cohort. Samples were collected at gestational weeks 8-11 (1st trimester placentas), 29-34 (early onset PE and control pregnancies) or 36-40 (healthy term pregnancies). Detailed description of the FINNPEC cohort has been reported previously [105].

Protein level measurements

Quantification of SALSA in amniotic fluid by ELISA
To quantify the levels of SALSA in protein extracts (AF, meconium and feces) and in non-treated AF samples an enzyme linked immuno-sorbent assay (ELISA) was set up. The samples were diluted in TBS/Ca\(^{2+}\) and coated onto Maxisorp plates (Nunc, Roskilde, Denmark). SALSA purified from saliva was used as a protein concentration standard.
### Materials and methods

#### Table 6: Proteins and antibodies used in this thesis

<table>
<thead>
<tr>
<th>Material</th>
<th>Origin</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombinant SALSA (rSALSA)</td>
<td>Expressed in CHO cells and purified by bacterial binding and EDTA-elution [152]</td>
<td>I</td>
</tr>
<tr>
<td>Saliva-SALSA</td>
<td>Purified by bacterial binding and EDTA-elution [152]</td>
<td>I, II</td>
</tr>
<tr>
<td>AF-SALSA</td>
<td>Purified by bacterial binding and EDTA-elution [152]</td>
<td>II, III</td>
</tr>
<tr>
<td>Recombinant MBL (rMBL)</td>
<td>Expressed according to previously described [75]</td>
<td>I</td>
</tr>
<tr>
<td>Recombinant M-ficolin (rM-ficolin)</td>
<td>Expressed according to previously described [212]</td>
<td>I</td>
</tr>
<tr>
<td>Plasma purified L-ficolin</td>
<td>Purified according to previously described [217]</td>
<td>I</td>
</tr>
<tr>
<td>Plasma purified H-ficolin</td>
<td>Purified according to previously described [115]</td>
<td>I</td>
</tr>
<tr>
<td>Recombinant MASP-2 (rMASP-2)</td>
<td>Expressed according to previously described [193]</td>
<td>I</td>
</tr>
<tr>
<td>Plasma purified C4</td>
<td>Purified according to previously described [38]</td>
<td>I</td>
</tr>
<tr>
<td>Plasma purified C3</td>
<td>Purified according to previously described [186]</td>
<td>I</td>
</tr>
<tr>
<td>C1q</td>
<td>Quidel, USA</td>
<td>I, II</td>
</tr>
<tr>
<td>IgA</td>
<td>Sigma-aldrich</td>
<td>II</td>
</tr>
<tr>
<td>Plasma purified fibronectin</td>
<td>Chemicon, USA</td>
<td>III</td>
</tr>
<tr>
<td>mAb Mouse anti-SALSA (Hyb 213-06)</td>
<td>Bioperto, Denmark</td>
<td>I, II, III</td>
</tr>
<tr>
<td>mAb Mouse anti-M-ficolin (7G1)</td>
<td>Produced according to previously described [212]</td>
<td>I</td>
</tr>
<tr>
<td>mAb Mouse anti-MBL (Hyb 131-01)</td>
<td>Bioperto, Denmark</td>
<td>I</td>
</tr>
<tr>
<td>mAb Mouse anti-H-ficolin (4H5)</td>
<td>Hycult Biotechnology, The Netherlands</td>
<td>I</td>
</tr>
<tr>
<td>mAb Rat anti.MASP-2 (8B5)</td>
<td>Hycult Biotechnology</td>
<td>I</td>
</tr>
<tr>
<td>pAb Rabbit anti-C3c</td>
<td>Dako</td>
<td>I</td>
</tr>
</tbody>
</table>
Materials and methods

<table>
<thead>
<tr>
<th>Ab Description</th>
<th>Source</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAb Rabbit anti-C4c</td>
<td>Dako</td>
<td>I</td>
</tr>
<tr>
<td>pAb Rabbit anti-C1q</td>
<td>Dako</td>
<td>I, III</td>
</tr>
<tr>
<td>HRP rabbit anti-mouse IgG</td>
<td>Jackson ImmunoResearch Laboratories, USA</td>
<td>I, II, III</td>
</tr>
<tr>
<td>HRP goat anti-rabbit IgG</td>
<td>Jackson ImmunoResearch Laboratories</td>
<td>I</td>
</tr>
<tr>
<td>Alexa 488 goat anti-mouse IgG</td>
<td>Invitrogen, USA</td>
<td>I, III</td>
</tr>
<tr>
<td>Alexa 546 goat anti-mouse IgG</td>
<td>Invitrogen</td>
<td>III</td>
</tr>
<tr>
<td>Alexa 488 goat anti-rabbit IgG</td>
<td>Invitrogen</td>
<td>I, III</td>
</tr>
</tbody>
</table>

Table 7: Buffers used in this thesis

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Content</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coating buffer</td>
<td>15 mM Na$_2$CO$_3$, 35 mM NaHCO$_3$, pH 9.6</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Phosphate-buffered saline (PBS)</td>
<td>137 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 18.8 mM KH$_2$PO$_4$, pH 7.4</td>
<td>I, II, III</td>
</tr>
<tr>
<td>Tris-buffered saline (TBS)</td>
<td>140 mM NaCl, 5 mM Tris, pH 7.4</td>
<td>I, II, III</td>
</tr>
<tr>
<td>TBS/Ca</td>
<td>140 mM NaCl, 5 mM Tris, 1 mM Ca$^{2+}$, pH 7.4</td>
<td>I, II, III</td>
</tr>
<tr>
<td>TBS/Ca/Tween</td>
<td>140 mM NaCl, 5 mM Tris, 1 mM Ca$^{2+}$, 0.05 % Tween-20, pH 7.4</td>
<td>I, II, III</td>
</tr>
<tr>
<td>TBS/EDTA/Tween</td>
<td>140 mM NaCl, 5 mM Tris, 10 mM EDTA, 0.05 % Tween-20, pH 7.4</td>
<td>I, II, III</td>
</tr>
<tr>
<td>TTSB buffer</td>
<td>0.5 mM NaCl, 20 mM Tris, 0.05% Tween 20, pH 7.4</td>
<td>III</td>
</tr>
<tr>
<td>Veronal buffered saline (VBS)</td>
<td>142 mM NaCl, 5.0 mM sodium barbital, pH 7.4</td>
<td>I</td>
</tr>
<tr>
<td>VBS/Ca</td>
<td>142 mM NaCl, 5.0 mM sodium barbital, 1 mM Ca$^{2+}$, pH 7.4</td>
<td>I</td>
</tr>
</tbody>
</table>
Materials and methods

After blocking and washing SALSA was detected using a monoclonal (mAb) anti-SALSA Hyb 213-06 and HRP-conjugated rabbit anti-mouse antibodies. For development 1,2-phenylenediamine (OPD) tablets (Dako, Denmark) were used. The resulting color was measured at an OD of 492 nm by an iEMS Reader MF (Labsystems, Espoo, Finland). Data points were obtained from a dilution series to ensure that the readings were within a linear range. The resulting protein level measurements were based on a minimum of three separate readings. In some cases the protein levels were correlated to the total protein amount of the sample, which was measured by NanoDrop (Thermo Scientific).

**LC-MS/MS Mass Spectrometry**

Relative peptide abundance was used to quantify the amount of either SALSA protein or specific peptide-containing regions of SALSA protein in protein extracts from AF, meconium and fecal samples. The samples were separated by gel electrophoresis and divided into four smaller regions. NanoLC and LTQ-Orbitrap-MS analyses, including quality checks and machine calibrations, were performed as described [106]. An in-house database based on protein sequences expected to be present in the infant gut was used for MS/MS spectral identifications.
Materials and methods

Protein visualization assays

Western blotting
To detect SALSA by Western blotting, samples were diluted in and mixed with non-reducing SDS-PAGE loading buffer. Thereafter, the samples were run into a 4-12 % gradient SDS-PAGE gel (Life Technologies) and the proteins were transferred onto a nitrocellulose membrane (Life Technologies). After blocking, SALSA was detected using anti-SALSA (Hyb 213-06) and HRP-conjugated rabbit anti-mouse IgG antibodies. The bands were visualized by electrochemiluminescence.

Glycoprotein and lectin staining
Purified SALSA was run into an SDS-PAGE gel. The resulting gel was split in three parts and stained with silver nitrate or Periodic-acid Schiff reagent (Glycoprotein Staining Kit, Pierce). In addition a part of the gel was blotted onto a PVDF-membrane (Amersham) for staining with DIG-labelled sialic acid-specific Sambucus Nigra lectin (DIG Glycan Differentiation Kit, Boehringer Mannheim). After blocking the membrane was incubated with the lectin. Lectin binding was detected with anti-DIG-AP according to manufacturer’s instructions.

Immunohistochemistry
Paraffin embedded tissue sections (5 μm) were subjected to standard de-paraffination followed by antigen retrieval treatment. The kit UltraVision LP Large Volume Detection System (HRP Polymer Ready-To-Use, ThermoFisher
Materials and methods

Scientific) was used. mAb anti-SALSA was added (10 µg/ml) in antibody diluent (Dako). The sections were then subjected to incubations with primary antibody enhancer, HRP Polymer, 3-Amino-9-Ethylcarbazole (both from ThermoFisher Scientific) and finally counterstained by Mayer's hematoxylin and eosin. The sections were treated with ammonium and mounted using Aquatex (Merck, Germany).

For fluorescence immunohistochemistry paraffin embedded sections were prepared as above. Frozen sections (5 µm) were prepared by cryosectioning of freshly frozen samples and blocked with bovine serum albumin (BSA). Anti-SALSA was diluted to 10 µg/ml and incubated with the samples. For co-localization studies rabbit-anti cellular fibronectin (ab299, Abcam, UK) and rabbit anti-C1q (Dako) antibodies were used 1:1000 in Dako antibody diluent. Alexa 488-labeled goat anti-rabbit and Alexa 546-labeled goat anti-mouse antibodies (Invitrogen,) were used diluted 1:300 in PBS. When ex vivo SALSA binding was tested, an overlay was performed with non-diluted AF.

Protein interaction assays

ELISA binding assays

Binding of SALSA to a range of endogenous ligands was tested in an ELISA set up. Binding was tested to recombinant MBL (rMBL), rM-ficolin, L-ficolin, H-ficolin, C1q, C4, C3, IgA and fibronectin. The proteins were diluted in a coating buffer into concentrations varying between 1 and 10 µg/ml. After
Materials and methods

coating, the wells were washed with TBS/Tween containing either 1 mM Ca$^{2+}$ or 10 mM EDTA. For fibronectin the TTSB buffer was used for washing. SALSA protein was added at 0.5-1 µg/ml. The Ca$^{2+}$-dependency of the binding was investigated by adding 10 mM EDTA and omitting Ca$^{2+}$ from the buffer (fibronectin only). After incubation binding was detected using anti-SALSA Hyb 213-06 and HRP-conjugated rabbit anti-mouse antibodies. The color reaction was developed as described above.

**ELISA competition assays**

Competition of binding between SALSA, MBL, MASP2 and carbohydrate ligands was tested by ELISA. In one assay rSALSA (0.1 µg/ml) was coated on microtiter plates. rMBL (1 µg/ml) was mixed in the fluid phase with mannose, GlcNAc or glucose (all from Sigma) in concentrations ranging between 0-100 mM in TBS/Ca$^{2+}$. The samples were incubated in the SALSA-coated wells. In another assay mannan (10 µg/ml) was coated on the plate. rMBL (0.5 µg/ml) was mixed with rMASP-2 (0.1 µg/ml) in TBS/Ca$^{2+}$ and rSALSA was added in final concentrations ranging between 0-1.5 µg/ml. The samples were then incubated on the plate. For both assays binding was detected with anti-MBL and/or anti-MASP2 and HRP-conjugated rabbit anti-mouse IgG antibodies. The color reaction was developed as described above.
Complement assays

Measurement of complement activation by SALSA in solution
The effect of fluid-phase SALSA on C activation was tested using the Wieslab® Complement System Screen ELISA assay (Euro Diagnostica, Sweden). SALSA (0-10 µg/ml) was diluted in normal human serum (NHS) and added to ELISA wells coated with specific activators for the three different C pathways. Activation of C was measured as generation of the C5b-9 complex onto the activating surfaces according to the manufacturer’s instructions.

Measurement of complement activation by surface-coated SALSA
In an ELISA assay mannan (10 µg/ml) or rSALSA (0.5 µg/ml) were coated on Maxisorp plates as described above. NHS, MBL-deficient serum, MgEGTA-serum and heat-inactivated serum (HIS) were diluted 1:10 and added. C4 and C3 deposition was detected by incubation with polyclonal (pAb) anti-C4c and C3c antibodies (Dako), followed by HRP-conjugated goat anti-rabbit antibody. The enzyme reaction was developed as described above.

Effect of SALSA on complement activation by C. albicans
The effect of SALSA on C activation by the yeast C. albicans was measured in a flow cytometry assay. C. albicans was a clinical blood culture isolate from the Helsinki University
Hospital laboratory (HUSLAB), identified using routine microbiological techniques. *C. albicans* was grown in yeast-extract peptone dextrose medium overnight at 30°C with shaking, washed and resuspended to $5 \times 10^7$ cells/ml. 100 µl of this dilution was used for each sample. rSALSA (0 – 1.5 µg/ml) was diluted in 10 % NHS, MBL-deficient serum, MgEGTA-serum or HIS and incubated with *C. albicans* for 30 min at 37°C. C4b and C3b deposition was measured using anti-C4c and anti-C3c antibodies followed by detection using Alexa 488-conjugated goat-anti rabbit IgG antibody. The yeast cells were fixed in 1 % paraformaldehyde and analyzed by CyAn ADP (Dako). Forward and sideward scatters were used to define the cell population and 10 000 events were routinely counted. The mean fluorescence intensity (MFI) values were used for quantification of the data.

**Bacterial binding assays**

**Bacterial culturing**

Group A streptococcus (GAS; ATCC 19615), group B streptococcus (GBS), ATCC T15508 and two clinical blood isolates, identified at HUSLAB and *S. gordonii*, DL1 Challis (20), were grown in Todd-Hewitt media O/N at 37°C. *E. coli* (urine isolate) and *Salmonella* serovar Typhimurium (fecal isolate) were grown O/N at 37°C with shaking in Luria broth.

**Binding of SALSA to bacteria**

SALSA binding to GBS was studied in a flow cytometry assay. GBS was grown as described above and resuspended to $1 \times 10^6$ cells/ml. Volumes of 100 µl of these dilutions were used
Materials and methods

for each sample. AF-purified SALSA (0-3 μg/ml) was incubated with the microbes followed by washing with VBS/Ca²⁺. SALSA binding was detected using mAb anti-SALSA (Hyb 213-06) and Alexa 488-coupled goat-anti mouse IgG antibodies. The microbes were fixed in 1 % paraformaldehyde and analyzed by CyAn ADP as described above.

Bacterial binding of SALSA from biological fluids was analyzed in a Western blotting-based assay. AF, meconium or fecal protein extracts were diluted to a final SALSA concentration of 0.5 μg/ml and incubated with 10⁹ bacterial cells. After centrifugation (10 000 g) the supernatants and pellets were collected. The bacteria were incubated in 50 μl non-reducing SDS-PAGE loading buffer (Life Technologies) containing 10 mM EDTA. Using Western blotting, SALSA in the original solution was compared to SALSA in the supernatants after absorption with bacteria and after treatment with 10 mM EDTA.

**SALSA-mediated inhibition of MBL binding to microorganisms**

The effect of SALSA on the binding of MBL to *C. albicans* and *E. coli* was studied in a flow cytometry assay. *C. albicans* and *E. coli* were grown as described above. *C. albicans* was resuspended to 5 × 10⁷ cells/ml and *E. coli* to 2.4 × 10⁸ cells/ml. Volumes of 100 μl of these dilutions were used for each sample. rMBL (0.9 μg/ml) was mixed with rSALSA (0-4.5 μg/ml) and incubated with the microbes. After washing
Materials and methods

with VBS/Ca\textsuperscript{2+} MBL binding was detected using an anti-MBL antibody and Alexa 488-coupled goat-anti mouse IgG antibody. The microbes were fixed in 1 % paraformaldehyde and analyzed by CyAn ADP as described above.

Coagulation assays

Effect of soluble SALSA on coagulation
Basic coagulation assays such as Thrombin Time and Activated Prothrombin Time measurements were performed using a coagulometer as described [8]. For thrombin time measurements 100 μl BC Thrombin reagent (Siemens, Germany) was added to 40 μl citrated plasma (at 37°C). For activated prothrombin time measurements 50 μl Dade Actin FSL reagent (Siemens) was mixed with 50 μl citrated plasma (at 37°C). After a 3 minute incubation 50 μl of 25 mM CaCl\textsubscript{2} was added to initiate coagulation. For both assays, SALSA was mixed with plasma in the fluid phase prior to initiation of coagulation at concentrations of 0 – 5 μg/ml.

Coagulation in the presence of surface-coated SALSA
The effect of surface-coated SALSA on coagulation was tested in an assay modified from the protocol published by Rose and Babensee [167]. SALSA (1 μg/ml) was coated on a Maxisorp plate as described above. Citrated plasma (100 μl, at 37 °C) and BC Thrombin reagent (100 μl) were added whereby coagulation was initiated. OD405 measurements were made at 20s intervals for 30 min using a FLUOstar
Materials and methods

optima reader (BMG Labtech, Germany). An increase in absorbance corresponded to the formation of the clot.

**Statistical analysis**

Student's paired, two-tailed *t*-test was used to calculate statistical significance of differences when comparing numerical values of SALSA protein levels, complement activation and coagulation. SALSA levels in AF samples from various disease groups were related to a list of clinical features. For this both a Pearson product-moment correlation test and a Spearman’s rank correlation test were performed.
Results and Discussion

This study has mainly focused on understanding the function of the SALSA protein, especially in the context of early life. We have analyzed certain expression patterns, identified novel ligands and found functional relevance in immunological systems, such as the complement system. An overview of the main results is given in Table 8.

Table 8: Main results obtained in this thesis

<table>
<thead>
<tr>
<th>Finding</th>
<th>Article</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secreted SALSA is found abundantly in amniotic fluid, meconium and feces</td>
<td>II, III</td>
</tr>
<tr>
<td>SALSA levels in amniotic fluid increase during pregnancy</td>
<td>III</td>
</tr>
<tr>
<td>SALSA is found in placental syncytiotrophoblasts and fibrinoids</td>
<td>III</td>
</tr>
<tr>
<td>SALSA is found in decidual endothelium</td>
<td>III</td>
</tr>
<tr>
<td>SALSA binds to MBL, ficolins and fibronectin</td>
<td>I, III</td>
</tr>
<tr>
<td>SALSA regulates complement activation</td>
<td>I</td>
</tr>
<tr>
<td>Variations in isoforms and glycosylations may affect SALSA function</td>
<td>II</td>
</tr>
</tbody>
</table>
SALSA in body fluids

SALSA in amniotic fluid (II, III)
This project was initiated by performing a screen of several body fluids for the presence of SALSA and led to the discovery of SALSA expression in amniotic fluid. This adds a new dimension to the potential functions of SALSA, which has earlier primarily been described in respiratory secretions (Table 1). To study the novel finding in greater detail we collected AF samples from women with normal healthy pregnancies undergoing routine screening or at term and investigated SALSA protein expression by Western blotting, ELISA and mass spectrometry.

The analysis of SALSA in AF revealed great individual variations. In Western blotting we observed size variations of the SALSA-positive bands in different AF samples (II, Figure 1). Furthermore, SALSA in several samples appeared as a single band and in others as a double band. SALSA presented either as very distinct bands or as smear of varying sizes on the blot. The AF samples were stained with Periodic-Acid Schiff and the sialic acid-specific Sambucus nigra lectin to visualize glycosylation. We observed a positive response indicating that SALSA in AF, like in other tissues, is glycosylated and contains carbohydrates with terminal sialic acids.

Size polymorphisms of SALSA have been described before and used to classify individuals into four groups \[^{46,103}\]. This grouping was shown to correlate with Lewis-antigen
expression, secretor status (expression of the α1-2fucosyl-transferase) and bacterial binding properties \textsuperscript{[46]}. The individual variations in SALSA observed in AF support the findings from saliva and lung aspirations and may indeed reflect such grouping in AF as well. In the future it will be of interest to compare purified AF-SALSA from individuals to search for a similar grouping and furthermore analyze if such grouping has a functional relevance for the SALSA protein found in AF. A thorough analysis would require larger volumes of the individual samples than we obtained in this study (only app. 10 ml per sample).

Individual variation was also found in the protein amount. In the ELISA analysis the concentrations of SALSA at term in the Spanish cohort ranged between 0-11.5 μg/ml (n = 14, mean: 2.1 ± 3.7 μg/ml) (II, Figure 2). The finding of very similar SALSA concentrations in the Finnish cohort (n = 9, mean: 2.1 ± 1.6 μg/ml) (III, Figure 1) increases the validity of the measured levels. To further understand the expression of SALSA during the course of the pregnancy, samples were also collected at an earlier time point (before gestation week 20, n = 20). Here the SALSA concentration ranged between 0-1.8 μg/ml (mean: 0.7 ± 0.5). The general protein level in AF is known to vary during the course of pregnancy and the SALSA concentrations were therefore related to the total protein amount. This led to the finding that the relative SALSA levels doubled from 0.15 ‰ to 0.3 ‰ of the total protein amount from the first to the third trimester (p<0.05) (III, Figure 1A).
Results and discussion

Amniotic fluid SALSA levels in normal and complicated pregnancies (II, III)

Dysregulated immunological responses have been suspected to be involved in the pathogenesis of certain pregnancy complications. Therefore we investigated the levels of SALSA in AF from patients with PE, IUGR, GDM or DM diagnosed prior to the pregnancy. Included in the analysis were additional samples taken before 20 weeks of gestation from patients who later developed PE. The concentrations of SALSA were related to the total protein amount in the samples and compared to AF samples from healthy age-matched controls.

At the early stage of pregnancy the concentrations of SALSA were 0.7 ± 0.5 μg/ml (n = 18) for the control group and 1.4 ± 1.4 μg/ml (n = 9) for the PE group (p = 0.09). When the values were related to the total protein levels the values were 0.17 ± 0.18 %₀ for controls and 0.29 ± 0.29 %₀ for PE (III, Figure 1B). Possibly because of the small number of samples, these differences were not statistically significant. Nevertheless, they suggest a trend of increased levels of SALSA at an early time point of patients who later develop PE and warrant studies in a larger population. Predictors of PE would be desperately needed. When we compared SALSA levels in the various pregnancy groups close to term, no significant differences were seen (III, Figure 1C). We also tested association between certain size polymorphisms and disease groups by Western blotting, but found no correlation. In both PE groups (early and at term) both high and low levels of SALSA were observed. This inter-group
Results and discussion

variation still lacks an explanation, but could indicate the existence of different disease subgroups with different pathological mechanisms. We tested if the SALSA levels correlated to the severity of the disease. However, no such link was found.

AF is a bioactive medium with constituents actively secreted by cells lining the amniotic cavity and, during the early stages, liquid filtered from the maternal blood \cite{20}. As gestation progresses, AF includes a significant volume of fetal urine (up to 50\%) \cite{20}. We tested adult urine but did not detect SALSA protein expression. The origin and thus the specific function of SALSA in AF is not yet clear. In this work we found several indications of placental production of SALSA (III, Figure 2). However, to fully determine the origin of SALSA in AF, the glycosylations could be studied in detail. Differences in e.g. the secretor status between the mother and the child would help to specify which tissue (maternal or fetal) produced the protein. Despite the fact that there can be differences in glycosylation even within the same individual, certain Lewis antigens found on SALSA in the AF of a Se(-) mother, would directly tell that SALSA is produced by fetal tissue.

The function of SALSA has been linked to both innate immunity and epithelial differentiation \cite{111,121}. Both types of functions could be mediated by SALSA in AF. In mice, a clear role of SALSA in epithelial differentiation from trophoderm to fully differentiated cells, including specific interactions with integrins and galectin-3 has been
Results and discussion

described [3]. SALSA knock-out proved to be lethal for the developing fetus in a mouse model [188]. In addition, seeding of mouse embryonic stem cells on a matrix of the rabbit SALSA ortholog, hensin, induced development of columnar epithelia [188]. This suggests that the developing cells in the human fetus could utilize SALSA in a similar way. The developing fetus could thus be the source of SALSA in the AF. The function would then be linked to epithelial cell differentiation.

In contrast, if the SALSA levels are indeed elevated in early pregnancy of women who later develop PE, this will most likely be a reflection of what happens at the feto-maternal interface in the placenta. An unbalanced meeting of fetal and maternal tissue could affect production of SALSA. This would support the hypothesis that SALSA is produced in the placenta and transported into the AF. This type of secretion is more suggestive of an immunoprotective role towards the fetus, perhaps helping to keep the AF sterile by the antimicrobial functions of SALSA. Thus we do not expect that elevated levels of SALSA in AF would be directly involved in the etiology of PE.

Intriguingly, the relative levels of SALSA increase towards parturition. This suggests that the role of SALSA becomes increasingly important. From an immunological perspective, it would be relevant to coat the fetus in anti-microbial effector molecules, such as SALSA, before it leaves the sterile amniotic cavity and encounters the first microorganisms of the maternal vagina and the surrounding environment.
Results and discussion

Likewise, an increase in SALSA levels would match the increase of cells in the developing fetus. Towards parturition SALSA could be produced mainly by the fetal mucosal tissues, which are constantly secreting molecules and fluids into the AF [20].

Studies have shown that SALSA levels correlate with lung maturity. Premature newborns had lower SALSA levels in tracheal aspirates right after birth compared to full term babies [128]. However, the same study showed that neonatal infection led to increased lung SALSA levels at birth [128]. Preterm infections are a known cause of premature birth and fetal death [52]. Interestingly, the surfactant proteins SpA and SpD, known ligands of SALSA, have been described to have a direct effect on the outcome of pregnancy in mouse models. Mice overexpressing rat SpA or rat SpD were found to have increased AF levels of the immunosuppressive cytokine IL-10 and the inflammatory marker TNF-α after LPS-induced preterm birth. Furthermore, preterm birth was induced with a lower dose of LPS in mice overexpressing rat SpD compared to controls [173,174]. Supporting this, a recent study found SpA/D double knock-out mice to have a delayed parturition, and an altered expression of several inflammatory markers in the myometrium [127]. It is possible that SALSA in AF interact with these molecules to exert their functions.

The adaptive immune system of the fetus and newborn is not fully developed. Therefore, they rely mostly on the innate immune system to defend against microbial invasion. At the
Results and discussion

same time, the interaction with the epithelium and the microbiota has been shown to be essential for the development of the epithelium [23]. Given the previously described functions of SALSA it is very likely that SALSA in AF plays a role linking these two effects. As such, it may be produced both by the epithelia of the developing fetus and by the amniotic epithelium. Solutions containing the protein would bathe all the mucosal surfaces and the developing epithelium and support the correct epithelial differentiation. Simultaneously, SALSA would, with its bacterial binding abilities, help to keep the AF sterile. In addition, the presence of the protein already from birth would allow SALSA to interact with the initial colonizing microbiota and possibly link the interaction between the microbiota and the developing neonatal epithelium.

**SALSA in the infant intestine (II)**

While studying AF, we also collected samples from other mucosal surfaces of the infants. From a total of nine individuals meconium was collected approximately 2 hours after birth and fecal samples were obtained one week after birth. Using Western blotting we identified SALSA in the samples, again exhibiting great individual variation. The protein levels were determined by ELISA and LC-MS/MS mass spectrometry (II, Figure 2). The protein concentration measurements obtained by ELISA gave average values of 45.8 μg/ml (range 2.8-294.6 μg/ml) for meconium and 22.4 μg/ml (range 0.1-62.9 μg/ml) for feces. The relative protein abundance of SALSA in the samples was obtained by mass
Results and discussion

spectrometry. Interestingly we found the levels to be 4.16 % for meconium and 2.81 % for feces. In some meconium samples SALSA was the most abundant protein overall.

During fetal development AF enters the gastrointestinal tract and contributes to the composition of meconium [175]. The gastrointestinal tract undergoes rapid growth from the fifth week of gestation through birth. During this phase AF flows through the digestive tract. Studies of infants with intestinal atresias and animal models have revealed that AF has both a physical and a trophic effect on the intestinal maturation [15, 175]. Thus, even during the later stages, the presence of SALSA at the fetal mucosal surfaces could be important for the development.

A key feature of the innate defense system is to accommodate the change from the sterile fetal life to the encounter with the first colonizing microbiota at a time when the adaptive immune system is still immature. Given the bacterial binding abilities of SALSA, and the finding that it is one of the most abundant proteins in the infant gut, SALSA may play an important role in the early selection of the colonizing microbes. Immunological molecules are present in the breast milk provided by the mother [24]. Evidence for the presence of SALSA in human breast milk has proven to be controversial. In one report no SALSA was detected, whereas a more recent study found high amounts of SALSA in breast milk [34, 165]. Most likely the discrepancy arises from individual variations, which we have also observed in our samples. Indeed, if SALSA is present in
Results and discussion

breast milk, it should also show up in the fecal samples. A preliminary study following SALSA in the gut of 4 infants in the first 100 days of life showed a difference in SALSA levels between breastfed (n = 2) and formula-fed (n = 2) infants (manuscript in preparation). This supports the hypothesis that maternal breast milk contributes to the SALSA-content in the intestines of the newborn. Regardless of the origin, it appears that SALSA is an important molecule at the mucosal surfaces in early life.

Comparison of SALSA in amniotic fluid, meconium and feces (II)

To understand the expression of SALSA at different mucosal surfaces, we compared the levels and the peptide composition of the protein in AF, meconium and feces. Just like we had observed individual differences in the presentation of SALSA in Western blotting from different individuals, we also observed differences between SALSA from different body compartments, even for the same individual. The protein abundance was 0.53 % for AF, 4.16 % for meconium and 2.81 % for feces (II, Figure 2). Although the largest pool of SALSA is found in the gut after birth, AF still constitutes one of the largest physiological reservoirs of SALSA (~1.5 mg) because of its large volume during late pregnancy (~700 ml at birth) [20, 66, 121]. So far the scientific community has primarily used SALSA purified from saliva as a source for experimental work. In the future AF may be a more easily accessible source of the physiological and functionally active form of the protein. Using AF rather than
Results and discussion

Saliva will also allow purification of higher amounts of SALSA protein from a single person. This will make it easier to analyze the individual differences in the protein.

As described in the literature review, several isoforms and glycoforms of SALSA have been described in tissues. However, SALSA has never before been compared in different mucosal compartments from the same person. It would have added great value to include maternal and fetal saliva samples in this study. However, unfortunately, the sampling protocol was decided before the SALSA-studies were planned, and we were not able to collect the saliva samples. Similarly, analysis of correlating breast milk samples would have been interesting for this study. These were not collected for the same reasons. In this study a clear difference between the proteins found in AF, meconium and feces were observed. Overall, SALSA in meconium appeared with a smaller molecular weight than that in AF. The same seemed to be the case for fecal SALSA. However, the difference to AF SALSA was not so clear-cut (II, Figure 1). A possible explanation for this difference could be that AF SALSA is more extensively glycosylated than the intestinal form of the protein. In other tissues polymorphisms have been ascribed to differences in both glycosylations and the protein backbone \[46, 66, 98, 120, 124\]. ABO and sialyl-Le\[^a\] carbohydrate antigens have been found to vary on SALSA found in saliva, respiratory secretions and tear fluid \[46, 180\]. However, in these studies the samples were not collected from the same individuals. It is therefore difficult to conclude if the observed differences originate from
Results and discussion

individual glycosylation differences or from tissue-specific differences.

The proteomics data obtained allowed us to analyze in detail the polypeptide composition of SALSA molecules from different sources. This revealed a difference in the abundance of certain peptides within the expected full-length transcript of SALSA (II, Figure 3). Some peptides (found in the ZP domain) were relatively more abundant in the AF samples, whereas some peptides (found in the SRCR domains) appeared comparatively more abundant in the intestinal samples. This finding suggests that in addition to the above described glycosylation differences, the size polymorphisms observed in the different tissues are also a result of differences in the protein backbone. Thus, we conclude that the size polymorphisms of SALSA arise from differential protein glycosylations and peptide-composition variations between AF, meconium and fecal SALSA, even within a single individual.

By comparison to crystal structures of domains found for homologs of SALSA we could link the differences in the peptide sequences to the potential established functions of SALSA (II, Figure 4). Interestingly, we discovered that the peptide sequence more often found in intestinal samples overlapped with the bacterial binding sequence of the SRCR domains, RVEVLYxxxSW. The peptide that was more abundant in the AF samples was found within a loop of the ZP domain believed to be responsible for protein dimerization. Our analysis of the relative abundance of
certain peptide sequences thus suggests the existence of structural differences between AF-derived and intestine (meconium and fecal) -derived SALSA in regions with a direct link to suggested protein functions, i.e. binding to bacteria via the SRCR domain and extracellular communication/polymerization via the ZP domain.

The mechanism of epithelial differentiation by the SALSA ortholog hensin has been described in detail in the rabbit [4]. An essential factor for the function of SALSA was the polymerization of hensin mediated by galectin-3. Since then, human SALSA and galectin-3 have been shown to interact and co-localize in the intestine, thus supporting a similar function for SALSA in the human gut [169]. ZP domains are found in hundreds of extracellular matrix proteins. Often these proteins polymerize through the ZP domain into fibrils or matrices and aid in transforming the cell shape and in creating polarization [150]. The ZP domain is therefore a potential site for the polymerization interactions when SALSA is deposited into the ECM.

Differences in specific expression patterns of SALSA, showing varying numbers of SRCR domains, have been described. However, all mRNA transcripts and proteins analyzed in various tissues by others, so far, have been found to display a ZP domain [120, 121, 151]. Our peptide results, showing a lower abundance of ZP domain-containing SALSA in the gut, do not necessarily mean that the ZP domain is not expressed here. It is possible, that the SALSA we find excreted with the feces after the passage through the
gastrointestinal tract has become modified. For example, the protein deposited into the ECM could have been cleaved by enzymes of the host or microbes. Structurally the ZP domain is found in the most C-terminal part of the protein, with the SRCR domains protruding from that. An electron microscopy picture revealed that polymerized SALSA is tied together in one end, with the rest of the protein extending into the surrounding space. Indeed, the utilization of the ZP-domain in polymerization and deposition into the ECM would shield it from the gut lumen, while the SRCR parts of the protein extend into the mucus layer. Here they would be accessible by colonizing bacteria but also degrading enzymes.

Ligand binding by SALSA

**Novel ligands of SALSA (I, II, III)**

This thesis work identified three new ligands of SALSA. In addition, we have confirmed previously described interactions. We have shown binding of purified or recombinant SALSA to MBL, M-ficolin and fibronectin (I, Figure 2 and III, Figure 7). All interactions were calcium-dependent. The previously described binding to C1q was confirmed. Binding was also observed to H- and L-ficolin, however with an apparently weaker affinity. In our assays we coated polystyrene plates with similar concentrations (w/w) of complement proteins and detected the bound SALSA with the same anti-SALSA antibody. The use of the same antibody allowed us to compare the signal from one protein to another. However, many of the complement
Results and discussion

proteins such as C1q, MBL and ficolins exist in various multimeric forms, making it difficult to assess the actual molar ratios of the interactions. In addition, the coating of these molecules on a plastic surface may interfere with the SALSA binding. Thus, it is possible that the binding of SALSA to H- and L-ficolin is as strong, or even stronger, than the MBL and C1q binding.

In addition to the binding assays using purified proteins, we also tested the ability of SALSA in AF and meconium to bind to IgA, MBL and C1q. Interestingly, SALSA in the intestine showed diminished ligand interactions. Only SALSA in AF retained the binding ability (II, Figure 6).

Simultaneously with our studies, other novel ligands for SALSA have been described by others. These include galectin-3, vascular endothelial growth factor, the growth factor Delta-like 4 and trefoil factor 3 \cite{112,130,164}. SALSA has so far had a great number of endogenous binding partners described. However, as mentioned above, it can be difficult to determine the strength and thus physiological relevance of these interactions. Microscale thermophoresis is a novel technique designed to measure protein interactions in the fluid phase \cite{87}. In the future it will be relevant to test separate Kd-values of the SALSA-ligand binding alone and in the presence of competing ligands. This will provide a more physiological view of the relevant SALSA binding partners.
SALSA-MBL interaction (I)

C1q was shown to bind SALSA through the globular domain in a region close to the immunoglobulin-ligand binding site \[^91]\). In order to investigate if MBL interacts with the glycosylated structures of SALSA, the MBL-SALSA interaction was compared to the binding of MBL to known carbohydrate ligands of its carbohydrate recognition domain (CRD). We coated SALSA on a plate in the presence of mannose, GlcNAc or glucose. Even at 100 mM concentrations we found no effect on the MBL binding (I, Figure 3B). However, when others performed a similar assay, they saw more than 60% inhibition of the SALSA-MBL interaction when 5 mM fucose was added to the fluid phase. They also observed inhibition with galactose and glucose \[^55]\). One explanation for the discrepancy of the effect of glucose could be that Gunput et al. only used SALSA purified from a single donor, whereas we used SALSA purified from a saliva pool. Based on the individual glycosylation patterns described above, our SALSA samples are more likely to contain several different carbohydrate structures. The inhibition of the protein binding by fucose and for some individuals also glucose lead us to conclude that these sugar moieties are important for the binding of MBL to SALSA.

We also performed a reversed assay. Here the effect of fluid-phase SALSA on the binding of the MBL-MASP2 complex to its carbohydrate ligands was tested (I, Figure 4). Addition of SALSA to the fluid phase interfered with the carbohydrate binding of the MBL-MASP2 complex. This suggests that SALSA binds directly to the carbohydrate binding site of
Results and discussion

MBL. Alternatively, SALSA may bind to the CRD in a fashion that disrupts the carbohydrate binding through steric hindrance. It was recently shown, that MBL binds to the \( \text{Le}^b \) antigen, a fucose containing oligosaccharide, which is also found on the surface of SALSA from secretors \([55]\). Together these findings strongly suggest that MBL binds via the CRD to the \( \text{Le}^b \) antigen of SALSA. However, in the presence of 100 \( \mu \text{g/ml} \) competing \( \text{Le}^b \), residual binding between SALSA and MBL was still observed, and thus other sugar structures on SALSA are most likely also involved \([55]\).

MBL belongs to a group of structurally similar proteins called collectins. SpA and SpD also belong to this group of proteins \([57]\). Furthermore, C1q shares a great deal of structural resemblance with the collectins. There may therefore also be similarities in how these proteins exert their functions in concert with SALSA. SALSA binds to the CRD of SpD, however to a site different from the carbohydrate-binding site \([208]\). The glycosylation of SALSA was directly implicated in this interaction \([61]\). The co-operative effects of SpA, SpD and SALSA on agglutination of microbes, viral neutralization as well as inhibition of hemagglutination were shown \([208]\). However, it was found that this effect was not dependent on the protein interactions. For some donors certain glycosylations of SALSA increased the affinity of SALSA towards SpD, which resulted in an inhibition of the anti-microbial effect. In contrast, other donors had a normal anti-microbial response \([61]\). As described above for MBL, the different outcomes of SALSA-ligand interactions are strongly
Results and discussion

dependent on the glycosylation of SALSA. In most of our assays we used pooled samples of SALSA for the protein interaction studies. However, it is possible that we could have found clearer differences in the protein interactions, if we had selected our samples, e.g. based on Se(+) and Se(-) status. When we used individual AF and meconium samples for ligand binding studies (II, Figure 6), we did observe both glycosylation differences and differences in protein interactions. However, the differences were related to the tissue-origin of the SALSA sample (AF vs. meconium) and not to variations between the different donors.

The interplay between SALSA and its many endogenous ligands is evidently complex. It appears that sugar structures on SALSA are often part of the ligand binding. This is observed especially for the structurally similar collectins and C1q. Depending on the specific phenotype of SALSA, this can have adverse outcomes on the anti-microbial activity. Taken together, it is obvious that a better understanding of the specific interactions of the various forms of SALSA with the many endogenous ligands will help to understand how the innate immune defense system orchestrates its anti-microbial responses at the mucosal surfaces.

Complement regulation by SALSA

Complement activation by surface-bound SALSA (I)
To see the effect of surface-coated SALSA on C activation rSALSA was coated on a microtiter plate, and C4 and C3 deposition was measured after incubation of 10 % NHS,
Results and discussion

MBL-deficient serum, MgEGTA-serum or HIS. A parallel experiment was done using a mannan-coated plate (I, Figure 6). We observed C activation on both SALSA- and mannan-coated plates. Complement activation was reduced when MBL-deficient serum was used, but not abolished. When MgEGTA-serum was used, C4 and C3 deposition to both SALSA and mannan was almost completely eliminated. We thus observed SALSA-mediated activation of C both through the classical and the lectin pathways of C.

Approximately 30% of the total C activation was based on the presence of MBL, confirming the relevance of our observed interaction between SALSA and MBL. However, the finding of residual C activation occurring even when MBL-deficient serum was used supports the previously described SALSA-mediated C activation through C1q and the classical pathway [17]. In addition, the interaction with the ficolins may also be a source of C4 and C3 deposition. As described, the primary expression of SALSA is found at the mucosal surfaces. Complement proteins are found mainly in serum. However, when damage occurs to the epithelium, tissue components such as C1q and MBL mix with components of the mucosal surfaces [16, 147]. It may be in these particular scenarios that SALSA and C work together in the defense against invading microbes.

Complement activation by fluid-phase SALSA (I)
To evaluate the effect of SALSA free in solution, the C regulating properties were studied in an ELISA assay
Results and discussion

(Wieslab). Here we found that the presence of SALSA in serum inhibited the lectin pathway mediated C deposition in a dose-dependent way. No effect was seen on the classical or the alternative activation pathways (I, Figure 1).

To further analyze this effect we used a flow cytometry assay to measure C4 and C3 deposition on a microbial surface that does not bind SALSA (I, Figure 7). *C. albicans* was used as a model organism in the presence of varying concentrations of SALSA. NHS, MBL-deficient serum, MgEGTA-serum or HIS (all at 10 %) were used. When using MBL-deficient serum, MgEGTA-serum or HIS little or no C deposition was found. When NHS was used, we observed C activation against *C. albicans*. The effect of adding varying concentrations of rSALSA showed a weak inhibition of the C4 deposition and a clear dose-dependent decrease of the C3 deposition. These data suggested that the presence of SALSA in the fluid phase inhibits the activation of MBL-mediated complement activation and subsequent deposition of complement factors C4 and C3 on *C. albicans*.

The effects of SALSA on the C system appear contradictory at first glance. However, the discrepancy between activation and inhibition can be explained by the localization of SALSA. We found surface-coated SALSA to activate the complement system, in line with others [16, 55, 95]. However, when SALSA was present in the fluid phase, we observed a weaker deposition of C components on the relevant targets, e.g. *C. albicans*. What we saw is most likely a result of SALSA binding to C1q and MBL in the fluid phase. The two
Results and discussion

outcomes of this can be an excessive fluid phase C activation, thus locally depleting the C components, or simply that initial binding of SALSA to C1q and MBL keeps them from binding to their targets. This would suggest that both the MBL-SALSA and the C1q-SALSA interactions are similar to the lactoferrin-SALSA interaction in the sense that a decreased bacterial binding is observed \cite{118,142}.

Recently, it has been shown that the activation of the lectin pathway by SALSA is dependent on the expression of specific Lewis antigens. SALSA activation of C from people with a Se(+) status was significantly higher than that of people with a Se(-) status \cite{55}. In the future it will be interesting to see if this activation is also observed when SALSA is present in the fluid phase.

**Binding of SALSA to bacteria**

**Bacterial binding by purified SALSA**

GBS are major neonate pathogens, leading to sepsis, pneumonia and meningitis \cite{82}. Due to the abundance of SALSA in early life we tested the binding of SALSA to GBS. Purified SALSA was incubated with GBS and binding was measured in a flow cytometry assay. This revealed a dose-dependent binding of SALSA to GBS (Figure 5). A recent study showed that surface-bound SALSA bound to the pili of different GBS strains, and that fluid-phase SALSA mediated agglutination \cite{25}. Given our findings of SALSA in the gut of newborns, and the binding of SALSA to several GBS strains, we hypothesize that SALSA plays a role in the innate defense
Results and discussion

Figure 5: SALSA binding to GBS.

Flow cytometry analysis measuring SALSA binding to GBS. GBS (10^5 cells in 100 µl) was incubated with varying concentrations of AF-purified SALSA. Binding was detected using anti-SALSA (Hyb213-06) and Alexa-488 coupled rabbit anti-mouse IgG. Displayed are averages and SD’s of two experiments.

against GBS after birth. Following vaginal delivery or caesarian section, the newborn will be exposed differently to the maternal microflora. Hypothetically this effect on the microflora could influence the SALSA expression, however, we found no correlation between SALSA levels and a specific type of delivery.

GBS is a frequent colonizer of the female vaginal surfaces, and may also be a cause of neonatal infection [158]. In this
Results and discussion

situation, SALSA may also be part of the response against the invading pathogen. Like SALSA, MBL is also found in the AF \textsuperscript{[113]}. The physiological relevance of the SALSA-mediated inhibition of MBL-ligand binding was tested in a flow cytometry assay using \textit{C. albicans} and \textit{E. coli} (I, Figure 5). MBL bound to both strains, but SALSA did not. Recombinant SALSA and recombinant MBL were mixed and then incubated with the microbes. We found that SALSA mediated a dose-dependent inhibition of MBL binding to \textit{C. albicans} and \textit{E. coli}. It appears that the binding of SALSA to MBL directly interferes with the binding of MBL to the surface of the pathogen. This observation thus further supports that the inhibition of complement activation against \textit{C. albicans} is partly mediated by SALSA blocking the binding of MBL to its target, as suggested above.

IgA and lactoferrin are two endogenous ligands of SALSA with structures different from the collectins. Both IgA and lactoferrin were found to bind to the SRCRP2 peptide sequence of SALSA rather than to sugar structures \textsuperscript{[99, 118]}. Controversially, the SALSA-lactoferrin interaction inhibits the binding of SALSA to bacteria, whereas the SALSA-IgA interaction enhances the bacterial agglutination \textsuperscript{[99, 118, 172]}. Thus, our finding of SALSA inhibiting the anti-bacterial function of MBL adds to the list of protein interactions with potentially adverse functional outcomes. Although SALSA, IgA, lactoferrin and MBL are all anti-microbial proteins, they do not necessarily cooperate to enhance the anti-microbial response. One reason for this could be that the innate defense system has evolved a more fine-tuned response to
Results and discussion

pathogens. A full binding and downstream activation of C could lead to excessive immune response, in situations where it is not needed, e.g. on the mucosal surfaces.

The binding of the SALSA ligand SpD to influenza A virus has been shown to induce a strong respiratory burst response in neutrophils *in vitro*. This response was reduced by the addition of SALSA [209]. It has been suggested that this allows a regulated response by the neutrophils, with an increased uptake of IAV but without an excessive and potentially harmful burst response [111]. We suggest that the interactions between SALSA and MBL may exert a similar regulation to suppress excessive C activation.

**Bacterial binding by SALSA in biological fluids (II)**
We tested the bacterial binding ability of native SALSA found in biological fluids. In a Western blot-based assay we tested binding to GAS, GBS, *S. gordonii*, *E. coli* (a urine isolate) and *S. Typhimurium* (a fecal isolate). SALSA from AF bound to GAS, GBS, *S. gordonii*, and the *E. coli* test strain. We observed no binding of SALSA from AF to *S. Typhimurium* (II, Figure 5A and E). The binding of SALSA from intestinal samples was also tested. In contrast to AF-SALSA we observed no bacterial binding by SALSA in meconium (II, Figure 5B and F). The SALSA found in fecal samples taken one week after birth showed individual differences in the bacterial binding ability. SALSA in one sample showed a clear binding to GAS and *S. gordonii*, and a weaker binding to GBS (II, Figure 5C).
Results and discussion

In contrast, SALSA in feces from another individual did not bind to any of the three streptococcal strains (II, Figure 5D).

When we measured SALSA binding from biological solutions, we used Western blotting instead of flow cytometry. The Western blotting assay is easy to use for a general screen of binding. However, the data outcome is only semiquantitative. In certain cases we observed a complete clearance of SALSA from the initial biological material after incubation with the microbes (II, Figure 5A). However, in other samples we saw some elution from the bacteria even though SALSA was not completely cleared from the starting material (II, Figure 5C). Although the conclusion is that SALSA in this sample binds to the bacteria, in this case GBS, it does suggest a weaker binding compared to the samples where SALSA was completely cleared. In these situations, flow cytometry would have provided more quantifiable data. However, even after protein extraction, the materials such as meconium and feces are difficult to work with. In our hands it was not feasible to utilize them in flow cytometry. Thus, the conclusions were based on data obtained by the Western blotting assays.

In line with the above described variations in SALSA interactions with endogenous molecules, and the adverse effect on microbial binding, it is not surprising that we observed differential binding between SALSA from AF and the intestine. SALSA was the most abundant protein in the meconium of some individuals (II, Figure 3). The high levels of SALSA in the intestines of the newborns indicate an important function for SALSA. From the moment we are
Results and discussion

born, the bacterial colonization of the gut starts, and a specific individual microbiome is selected. For this, several different mechanisms, e.g. the interplay with the innate immune defense system at the mucosal surfaces will be utilized \(^{197}\). Given the high expression of SALSA in the gut, the specific bacterial binding abilities of the protein must affect the colonizing bacteria. We observed individual differences in the phenotype of SALSA and an individual specific selection of bacterial ligands. It is therefore possible that SALSA found at the mucosal surfaces could be a part of the interaction between the colonizing bacteria and the host, and thus aid in the selection of a specific composition of microbes.

**SALSA at the feto-maternal interface**

**SALSA localization in placenta (III)**

After discovering SALSA in AF, we investigated its expression in the surrounding tissues. We found SALSA expression in both the placenta and in the maternal decidua. We analyzed the specific localization of the protein using immunohistochemistry on frozen and paraffin embedded sections of healthy term placenta, PE term placenta and 1st trimester placenta (III, Figures 2-4).

In the term placenta a distinct positive staining of SALSA was observed intracellularly in the syncytiotrophoblasts (III, Figure 2C and D). The cytotrophoblasts were not positive for SALSA. The syncytium of some villi stained more strongly than others, indicating that the expression of SALSA is
Results and discussion

inducible rather than constitutive. So far, the factors regulating the expression of SALSA here are not known. In addition, we observed abundant focal and distinct staining of SALSA in fibrinoid structures at various locations (Figure 6). Fibrinoids are divided into matrix-type fibrinoids, which are secretion products of extravillous trophoblasts, and fibrin-type fibrinoids, which are blood coagulation products with large amounts of fibrin [86]. SALSA was found primarily in fibrin-type fibrinoid. These SALSA-positive structures were found at the edges of the villous trees facing the maternal side. In addition, SALSA was found within individual necrotic villous structures and also in larger necrotic areas with massive fibrin formation. No major differences were observed in the staining pattern of SALSA between healthy and PE pregnancies.

Figure 6: SALSA in placental fibrinoid structures.

Frozen sections were stained with anti-SALSA antibody (Hyb 213-06) and Alexa 488-conjugated goat anti-mouse IgG. A) Fibrinoids are often found lining the maternal side of the villi. B) When the syncytiotrophoblast layer is disrupted, fibrinoid is deposited between the syncytium and basement membrane (white arrow). In conjunction with this fibrin-formations are often seen protruding into the intervillous space. Magnification: 400×.
Results and discussion

To our knowledge, the SALSA protein has not previously been described in the placenta. However, SALSA has been found expressed on cervical and vaginal epithelial cells in humans \[^{185}\]. Furthermore, in rodents SALSA was observed directly in the uterine epithelium and in both rodents and primates SALSA mRNA expression was markedly increased after estrogen stimulation \[^{198}\]. A specific sugar structure expressed on SALSA has been implicated in the actual implantation. It was found that the actual implantation is mediated through the interaction of L-selectin expressed on the surface of trophoblast cells and the MECA-79 carbohydrate epitope expressed on the uterine epithelium \[^{50}\].

Ligands of SALSA, C1q, MBL, SpA and SpD, were recently described to be expressed in the decidua and placenta of early and term pregnancy \[^{2,109,215}\]. C1q, SpD and SpA were found expressed by decidual stromal cells and invading trophoblasts of 1st trimester pregnancy \[^{2,109}\]. In term placenta MBL, SpA and SpD were also found in the syncytiotrophoblast layer \[^{215}\]. These SALSA ligands have all been suggested to be involved in the process of trophoblast invasion, embryo implantation and placental development \[^{2,109,215}\].

**SALSA in fibrinoids (III)**

The function of fibrinoid has so far been linked to adapting the intervillous space to the altering flow conditions, to control the growth of the sprouting villous trees and to
function as a substitute barrier wherever the continuity of the syncytiotrophoblast layer at the feto-maternal interface has broken down. Furthermore, it has been shown that the formation of fibrinoid is utilized by trophoblasts to re-epithelialize the villi. In the early stages of pregnancy the oxygen pressure in the placenta is very low. After 10-12 weeks of gestation, the intervillous space is flooded by maternal blood. This may cause oxidative stress with subsequent injury to the syncytium. After a local injury or damage the syncytiotrophoblast layer may be interrupted. We often found strong fibrinoid formation with SALSA deposition in conjunction with a disrupted syncytium. We also observed the fibrinoids separating the syncytiotrophoblast from the basement membrane and extending into the intervillous space. It appears that damage to the syncytium allows SALSA to enter the intervillous space and deposit into the formed fibrinoid. In some cases maternal blood will flow into the villus after a breach of the syncytium. When this had occurred we saw SALSA deposited in ring-like fibrinoid structures separating the syncytiotrophoblasts from the basement membrane (Figure 6B, white arrow).

Previous studies have indicated that SALSA and fibronectin are both involved in epithelial differentiation. SALSA and fibronectin may be involved in this process in the developing placenta, as well. Thus, SALSA secretion by the endothelial cells or by the damaged syncytium would result in the deposition of SALSA into the ECM, e.g. through interaction with fibronectin.
Results and discussion

The analysis of SALSA expression in the placenta showed that SALSA was deposited in fibrin-type fibrinoids, necrotic villi and irregularly on vascular endothelial cells. To identify further potential targets for SALSA an overlay of frozen placental sections with SALSA-containing AF was done (III, Figure 5). In addition to the previously described staining patterns, we observed binding of SALSA to a large part of the syncytial basement membranes and the endothelium of most capillaries and large vessels. This indicated that targets for SALSA are present in the endothelium either directly on the surface of the endothelial cells or in the extracellular matrix. We suspect that the deposition of SALSA in the tissue is more related to the availability of SALSA rather than the expression of specific targets because of e.g. tissue injury.

**SALSA in 1st trimester placenta**

In the sections from 1st trimester placenta fibrinoids were almost absent. Instead, we observed SALSA in the decidual endothelium of both small capillaries and larger blood vessels. The staining pattern was irregular suggesting that the expression of SALSA is induced under certain conditions. A recent study described for the first time that endothelial cells secrete SALSA into the ECM \[130\]. It was found that endothelium-derived SALSA bound galectin-3, affected Notch signalling and promoted proliferation, angiogenesis and vascular repair \[130\]. Our findings in the early human placenta and decidua provide further evidence that SALSA is expressed in blood vessels and is deposited into the ECM under physiological conditions.
Results and discussion

At the 1st trimester stage of a normal healthy placentation, angiogenesis is an ongoing process. As the pregnancy continues, various areas of the placenta may experience either local hypoxia or oxidative stress. The syncytiotrophoblast may be damaged by oxidative stress when the maternal blood flow into the placenta is first established [73, 74]. In cases of failing blood flow the placenta will experience hypoxia. In both cases this may present a “danger signal” and one of the responses could be the expression of SALSA. Indeed, recent work with a SALSA knock-out mouse model suggested that SALSA is part of the endothelial cell response to hypoxia, as the mice showed an impaired recovery from ischemic hindlimb injury [130]. Although we have not yet been able to address the inducing factors of SALSA expression, together the described findings support a physiologically relevant function of SALSA in endothelial remodelling during placentation. Whether the role of SALSA in fibrinoid formation and the ECM of syncytiotrophoblast is linked to its function in the ECM of endothelial cells remains to be investigated. Indeed SALSA may be interacting with different ligands at the two locations. It is evident that both vascular development and syncytium regeneration require cell proliferation and migration. This happens in interactions with the underlying ECM, where SALSA could have an important function.

Effect of SALSA on blood clotting (III)
The complement and the coagulation systems are closely linked. It has been shown that several proteins of the
Results and discussion

coaagulation system can activate C, e.g. proteases such as plasmin and thrombin can cleave C3 \[^{114}\]. SALSA has previously been suggested to interact with fibrin/fibrinogen, platelets and erythrocytes \[^{129}\]. Our finding of SALSA in fibrin-type fibrinoids led us to investigate the role of SALSA in the formation of blood clots. We performed basic coagulation assays such as Thrombin Time and Activated Prothrombin Time measurements with SALSA present in the fluid phase. In addition we analyzed clot formation on a surface coated with SALSA by measuring absorbance at 405 nm of citrated blood plasma after initiation of coagulation (III, Figure 6). In these experiments the increase in absorbance correlates to the formation of the clot, however we did not observe any effect of SALSA on the extent or speed of clot formation. Our assays were performed in plasma without platelets. Thus, our results are not contradictory to the finding that SALSA aggregates and activates platelets \[^{129}\]. Müller \textit{et al.} gave indications that the interaction of SALSA with fibrinogen was stronger than with fibrin. Under physiological conditions, this would suggest that SALSA exerts its function before fibrinogen is activated into fibrin. In light of this, it is somewhat surprising that we did not see any effect on the formation of the fibrin clot. For coagulation assays citrated plasma is used, in order to inhibit the untimely activation of clotting. Many functions of SALSA are known to depend on the presence of calcium, and this could be an explanation for the apparent lack of effect on clotting \[^{154}\]. We performed our assays with addition of calcium, enough to initiate coagulation. However, it may be that SALSA requires even higher concentrations of calcium.
Results and discussion

Still, based on the data from our assays it appears that SALSA is not involved in the process of fibrinoid formation. Instead it seems SALSA is deposited after fibrin deposits have been formed.

**Co-localization of SALSA with complement and fibronectin (III)**

We have described fibronectin and confirmed C1q as endogenous binding partners of SALSA (III Figure 7 and I Figure 2). Both proteins have been linked to fibrinoids [86, 105]. To investigate if SALSA interacts with these proteins in vivo, co-localization of the two proteins with SALSA in human placenta was studied. Both fibronectin and C1q were found associated to the same fibrinoid structures as SALSA (III, Figure 8). Although some co-localization was observed in the fibrinoids, in particular for fibronectin, most SALSA staining was seen in the inner part of the fibrinoid structures, while fibronectin and C1q were located at the edges of the structures (Figure 7). It thus appears that SALSA is directly incorporated into the fibrinoid matrix, and to some extent the same applies for fibronectin (Figure 7A). However, the staining of C1q shows that the protein encapsulates the SALSA positive areas, especially in necrotic villi (Figure 7B). Our findings are supported by other studies where it was found that C components and SALSA were deposited in the same necrotic amyloid depositions in the heart tissue. However, no direct co-localization was observed [131].
Results and discussion

Figure 7: Fibronectin, C1q and SALSA in placental fibrinoids.

Paraffin embedded placental tissue sections of healthy term placentas were stained with anti-fibronectin antibody and anti-C1q antibody (both green). SALSA was stained using anti-SALSA antibody (Hyb 213-06, red). A) SALSA and fibronectin are found in the same structures, with some degree of co-localization (white arrow). B) SALSA is observed staining a fibrinoid ghost-like structure of a necrotic villus. C1q is deposited as a coating on the edge of the SALSA-positive fibrinoid (white arrow). Magnification: 200×.

In pre-eclampsia and other complicated pregnancies, excessive placental hypoxia and damage of the syncytiotrophoblast commonly occurs. In these situations the expression of SALSA could be induced locally at the same time as the clot formation takes place. C1q is known to target apoptotic cells, cellular debris and the ECM. Complement activation thus has important housekeeping functions \[135\]. Interestingly, MBL and the SALSA ligands SpA and SpD have been linked to similar processes \[90, 134, 143\]. A recent study described differences in the deposition of C1q and the complement inhibitor C4bp in fibrinoid structures between healthy and PE placentas \[105\]. We suggest that in the case of local placental ischemia and tissue damage, the function of SALSA is to help contain the necrotic process and the excessive
formation of fibrinoid through interactions with fibrin and fibronectin and thereafter participate in the removal of the debris through interactions with C1q and C. Dysregulation of C at the feto-maternal interface has long been suspected to be part of the etiology of PE \cite{107,108}. A recent study found deposition of C4d on the syncytiotrophoblast layer in PE placentas, but not in healthy controls \cite{27}. We did not observe a difference in the pattern of SALSA staining in PE placentas compared to placentas from healthy pregnancies. However, as there is more syncytial damage in PE than normal placentas, a greater activation of C and deposition of SALSA is likely to occur in PE. Secretor status and Le\textsuperscript{b} expression was shown to affect the C activation mediated by SALSA. These specific carbohydrate-structures have also been linked to pregnancy disorders such as IUGR and recurrent spontaneous abortion \cite{31,51}. In the future it will be of interest to understand more thoroughly the interaction between SALSA and individual C components in the human placenta. Furthermore a better knowledge of variations in the SALSA genotypes and phenotypes and how they could lead to altered C interactions could help us to understand the effect of SALSA on the placental development in both healthy and complicated pregnancies.
Conclusion

The early work on the function of the SALSA protein has primarily focused on understanding the interplay between control of epithelial re-generation and immune defense at the mucosal surface. However, the results obtained in this work, together with observations made elsewhere during recent years point to a much broader field of action. It has long been known that SALSA is a molecule with the ability to bind many different endogenous protein ligands, as well as microbes. Lately even gametes and stem cells have been suggested as targets for SALSA [5, 111, 188, 192].

Table 1 lists sites of SALSA expression in human known prior to this study. In this thesis work we have described novel sites, tissues and secretions, of SALSA expression. We have shown e.g. that the SALSA protein is present in amniotic fluid, and very abundantly in meconium and in infant feces. In addition, we have described SALSA in the placenta and in the decidua.

When combined, the different studies show a clear profile of constitutive SALSA expression at the mucosal surfaces. However, the protein production or secretion is clearly inducible in several other tissues. Therefore, a key to understanding the multi-faceted role of SALSA is to understand site- and time-patterns of its expression. It is not only interesting, where the protein is found, but also under
which circumstances. Both inflammatory inducers as well as other stimuli are most likely involved in activating the expression of SALSA. In these situations, SALSA exerts other functions than simple scavenging for bacteria or tissue debris. In the future, a better understanding of the inducing factors of SALSA, e.g. in the placenta, is needed.

Previously, it has been shown that different forms of the SALSA protein exist in both saliva and tear fluid. In this study we further describe individual variations of the protein in amniotic fluids and in intestinal secretions. We found, for the first time, that for a given individual, the protein may exist in several forms in different tissues. By comparing the descriptions of the types of variations in other tissue compartments, the differences between the SALSA proteins that we have described, are most likely due to alternative splicing and posttranslational modifications such as proteolysis and differential glycosylations [46, 153, 180].

The different forms of the SALSA protein were linked to varying functional abilities. Lately, certain isoforms have even been associated with the development of chronic inflammatory conditions such as Crohn’s disease [37]. To fully understand the role of SALSA, a better knowledge of these differences becomes essential. We have shown that SALSA may regulate the complement system to different outcomes, depending on whether it is found in the solution or in the solid phase. To further complicate things, a recent study found that individual expression of the carbohydrate structures attached to Se(+) and Se(-) status had an effect on
the ability of SALSA to regulate the C system \cite{55}. The activation of the C system by SALSA has already been suggested to be mediated by an interaction with the sugar-binding globular domain of C1q, and it is therefore not surprising, that these interactions are altered by differences in the secretor status – and possibly also other variations in glycosylations. As such, we described a binding of AF-SALSA to both C1q and MBL, but no binding was observed for SALSA obtained from meconium. In the future, it is important to understand how the phenotype of SALSA, glycoforms as well as protein isoforms, relate to the function of the protein. And furthermore, knowledge is needed on the possible individual regulation of the expression of specific SALSA phenotypes. Can there be temporal, e.g. hormonally induced, alterations in the function of the SALSA protein, as has been suggested previously \cite{1, 153, 198}? Or is it simply a question of individual genetic variation leading to the expression of SALSA molecules with a certain palette of functions? At the mucosal surfaces this could indicate the existence of particular phenotypes of SALSA. These would inhibit colonization by certain bacterial species, but not by others. Depending on the phenotype, this would have different outcomes on the colonizing microbiota of an individual. Supporting this hypothesis a recent study showed a link between the evolution of the SALSA gene and S. mutans, a known target of SALSA found on the mucosal surfaces \cite{151}. Most likely, both genetic and environment-induced variations affect the final expression of the SALSA protein.
Conclusion

The current study suggests that the function of SALSA is linked to an innate immune response to very adverse challenges such as microbial infection and tissue injury, e.g. in the placenta. SALSA may thus have a dual role: pro-inflammatory implied by the C activating properties, or anti-inflammatory in situations, where injured structures need to be contained, as in the placental fibrinoids. In the intestine, bacterial infections as well as chronic inflammatory conditions are associated with increased SALSA expression. The future hope is to understand how addition of SALSA would either aid the immune defense in the battle against invading pathogens or alleviate chronic mucosal inflammatory conditions. Perhaps one day restoring tissue homeostasis could be achieved with a pill of SALSA?!
References


C1q-binding Sites of Human C1r and C1s: a refined three-dimensional model of the C1 complex of complement. *J. Biol. Chem.* 284: 19340-19348.


References


References


References


References


References


References


References


References


References


References


References


