

CEACAMs as novel receptors for  
*Helicobacter pylori*  
outer membrane protein HopQ

Dissertation

zum Erwerb des Doktorgrades der Naturwissenschaften  
an der Medizinischen Fakultät  
der Ludwig-Maximilians-Universität München

vorgelegt von  
Verena Königer  
aus Mainburg

2015



Aus dem Max von Pettenkofer-Institut für Hygiene und Medizinische  
Mikrobiologie

Lehrstuhl: Bakteriologie

der Ludwig-Maximilians-Universität München

Komm. Vorstand: Prof. Dr. Rainer Haas

Dissertation

zum Erwerb des Doktorgrades der Naturwissenschaften

an der Medizinischen Fakultät

der Ludwig-Maximilians-Universität zu München

vorgelegt von

Verena Königer

aus

Mainburg

Jahr

2015

Mit Genehmigung der Medizinischen Fakultät  
der Universität München

Betreuer:

Prof. Dr. Rainer Haas

---

Zweitgutachter(in):

Prof. Dr. Annette Nicke

---

Dekan:

Prof. Dr. med. dent. Reinhard Hickel

Tag der mündlichen Prüfung:

02.11.2015

---

Meinen Eltern & meinen Schwestern

Eidesstattliche Erklärung:

Ich versichere hiermit, dass ich die eingereichte Arbeit zum Erlangen des naturwissenschaftlichen Doktorgrades (Dr. rer. nat.) mit dem Titel „CEACAMs as novel receptors for *Helicobacter pylori* outer membrane protein HopQ“ selbständig und ohne unerlaubte Hilfe angefertigt habe. Ich habe mich dabei keiner anderen als der von mir angegebenen Quellen und Hilfen bedient. Ich erkläre des Weiteren, dass die hier vorgelegte Arbeit nicht in gleicher oder ähnlicher Form bei einer anderen Stelle zum Erlangen des naturwissenschaftlichen Doktorgrades eingereicht wurde.

---

Ort, Datum

---

Verena Königer

Die vorliegende Arbeit wurde in der Abteilung Bakteriologie am Max-von-Pettenkofer Institut für Hygiene und Medizinische Mikrobiologie der Ludwig-Maximilians-Universität München unter der Leitung von Prof. Dr. Rainer Haas angefertigt.

**Teile dieser Arbeit werden unter folgendem Titel veröffentlicht:**

**Koeniger V**, Holsten L, Loell E, Busch B, Zhao Q, Bonsor D, Breithaupt U, Roth A, Kengmo Tchoupa A, Smith S, Müller S, Sundberg E, Zimmermann W, Fischer W, Hauck C, Haas R. *Helicobacter pylori* exploits human CEACAMs for adherence and translocation of CagA (Manuskript in Revision)

Javaheri A, Kalali B, **Koeniger V**, Bach N, Tegtmeyer N, Asche I, Sieber S, Hauck C, Haas R, Busch D, Backert S, Klaile E, Slevogt H, Schmidt A, Singer B, Gerhard M. Engagement of CEACAM receptors by *Helicobacter pylori* modulates immune responses (Manuskript in Revision)

---

# Table of contents

---

<b>Zusammenfassung</b>	<b>1</b>
<b>Summary</b>	<b>3</b>
<b>Introduction</b>	<b>5</b>
<b>1 Immunoglobulin superfamily</b>	<b>5</b>
1.1 Composition and structure	5
1.2 Function	6
<b>2 Carcinoembryonic antigen gene family</b>	<b>7</b>
2.1 CEACAM family	7
2.2 Pregnancy-specific glycoproteins	14
2.3 CEACAMs in other species and evolution	16
2.4 CEACAM and cancer	19
2.5 CEACAM recognition by pathogens	20
<b>3 <i>Helicobacter pylori</i></b>	<b>22</b>
3.1 Discovery of <i>H. pylori</i>	22
3.2 <i>H. pylori</i> transmission and prevalence	22
3.3 <i>H. pylori</i> and disease	23
3.4 Major virulence factors	23
3.5 <i>H. pylori</i> adhesins	26
<b>4 Aims of this thesis</b>	<b>29</b>
<b>Results</b>	<b>30</b>
<b>1 <i>H. pylori</i> binds to CEACAM-N-GFP constructs</b>	<b>30</b>
1.1 Establishment of a bacterial pull-down assay	30
1.2 <i>H. pylori</i> strain P12 binds CEA1-N-GFP and CEA5-N-GFP	33
1.3 CEACAM-binding of <i>H. pylori</i> strains	35
1.4 Specificity of CEACAM-binding	38
<b>2 HopQ is the adhesin binding to CEACAM1 and CEACAM5</b>	<b>40</b>
2.1 Generation of markerfree <i>H. pylori</i> OMP mutants	40
2.2 <i>H. pylori</i> binding is HopQ dependent	44
2.3 HopQ variations in different <i>H. pylori</i> strains	46
2.3.1 Sequence analysis of HopQI and HopQII	46
2.3.2 HopQI production in strains	49

---

<b>3</b>	<b>CEACAM producing cell lines and CagA translocation</b>	<b>52</b>
3.1	Test for CEACAM production on cell lines	52
3.2	CEACAM producing HEK293 cells and CagA translocation	55
3.3	<i>In vitro</i> infections with P12 $\Delta$ <i>hopQ</i> mutant and the <i>hopQI</i> and <i>HopQII</i> complemented strain	57
3.4	Knockdown of CEACAM1, 5 and 6 in AGS cells	59
<b>Discussion</b>		<b>61</b>
<b>1</b>	<b>The capacity of bacteria for CEACAM-binding</b>	<b>61</b>
1.1	The bacterial pull-down assay – Strength and weakness	61
1.2	CEACAM-binding of <i>H. pylori</i> strains and the effect of animal passage	64
1.3	CEACAM-binding specificity	67
<b>2</b>	<b>The CEACAM-binding adhesin HopQ</b>	<b>69</b>
2.1	HopQ is the adhesin interacting with CEACAM1 and CEACAM5	69
2.2	Mode of HopQ binding to CEACAM5	71
2.3	HopQ production in <i>H. pylori</i> strains	72
<b>3</b>	<b>Impact of the bacterial CEACAM interaction on CagA translocation</b>	<b>75</b>
3.1	CEACAM production and CagA translocation capability	75
3.2	HopQ dependency of CagA translocation	77
<b>4</b>	<b>Outlook</b>	<b>79</b>
<b>Material</b>		<b>81</b>
<b>1</b>	<b>Cell lines</b>	<b>81</b>
<b>2</b>	<b>Bacterial strains</b>	<b>81</b>
2.1	<i>E. coli</i>	81
2.2	<i>H. pylori</i>	82
2.3	Other bacteria	82
<b>3</b>	<b>Plasmids</b>	<b>83</b>
<b>4</b>	<b>Oligonucleotides and siRNA</b>	<b>84</b>
<b>5</b>	<b>Commercially available kits</b>	<b>86</b>
<b>6</b>	<b>Antibodies</b>	<b>87</b>
<b>7</b>	<b>Bacterial culture medium</b>	<b>88</b>
<b>8</b>	<b>Buffers and solutions</b>	<b>88</b>
<b>9</b>	<b>Chemicals and enzymes</b>	<b>90</b>
<b>10</b>	<b>Consumables</b>	<b>90</b>
<b>11</b>	<b>Equipment and devices</b>	<b>90</b>

---

<b>Methods</b>	<b>92</b>
<b>1 Working with eukaryotic cells</b>	<b>92</b>
1.1 Maintenance of Cell lines	92
1.2 Cryoconservation	92
1.3 Cell Staining	92
1.4 Knockdown by RNA Interference	93
<b>2 Working with bacteria</b>	<b>94</b>
2.1 Cultivation of bacteria	94
2.1.1 <i>E. coli</i>	94
2.1.2 <i>H. pylori</i>	94
2.1.3 Other bacteria	94
2.2 Generation of chemically competent <i>E. coli</i> (Top 10)	95
2.3 Generation of <i>H. pylori</i> knockout mutants	95
2.4 Transformation of <i>H. pylori</i>	96
2.5 Electroporation of <i>H. pylori</i>	96
2.6 Bacterial pull-down assay	97
2.7 <i>In vitro</i> infection experiments with <i>H. pylori</i>	97
<b>3 DNA methods</b>	<b>98</b>
3.1 Isolation of DNA	98
3.1.1 Plasmids	98
3.1.2 Genomic DNA	98
3.2 DNA concentration and quality	98
3.3 DNA extraction from agarose gels and enzymatic reactions	99
3.4 Amplification of DNA fragments	99
3.4.1 PCR under standard conditions	99
3.4.2 Colony PCR (Sambrook et al, 1989)	101
3.4.3 RAPD PCR	101
3.5 Cloning	103
3.5.1 Restriction of DNA	103
3.5.2 Ligation	103
3.5.3 Transformation of chemically competent <i>E. coli</i>	104
3.6 Sequencing	104
<b>4 Protein methods</b>	<b>105</b>
4.1 Bacterial lysates	105
4.2 SDS PAGE	105
4.3 Coomassie staining	105

---

4.4	Western blotting	106
4.5	Detection of proteins	106
4.6	PVDF membrane stripping	107
<b>Appendix</b>		<b>108</b>
<b>1</b>	<b>Supplementary data</b>	<b>108</b>
1.1	Strain PMSS1	108
1.1.1	Sequencing of the <i>sabA</i> and <i>sabB</i> locus ( <i>hopQI</i> gene)	108
1.1.2	Sequencing of the <i>hopQ</i> locus ( <i>hopQII</i> gene)	109
1.2	Strain SS1	110
1.2.1	Sequencing of the <i>sabA</i> and <i>sabB</i> locus ( <i>hopQI</i> gene)	110
1.2.2	Sequencing of the <i>hopQ</i> locus ( <i>hopQII</i> gene)	111
1.3	Protein sequence HopQI (PMSS1/SS1)	112
1.4	Protein sequence HopQII (PMSS1/SS1)	112
<b>2</b>	<b>List of figures</b>	<b>113</b>
<b>3</b>	<b>List of tables</b>	<b>114</b>
<b>4</b>	<b>Abbreviations</b>	<b>115</b>
<b>Literature</b>		<b>117</b>
<b>Danksagung</b>		<b>131</b>

# Zusammenfassung

---

Die oberflächenexponierten Rezeptoren der CEACAM-Familie sind an grundlegenden Prozessen in der Zelle wie Zell-Zell-Adhäsion, Homöostase und der Regulation des Immunsystems beteiligt. Die CEACAMs werden bevorzugt von Viren und Bakterien genutzt, um über deren Bindung die Wirtszellen zu infizieren. CEACAM1, 5 und 6, die typischerweise auf Epithelzellen vorkommen, und CEACAM3, das nur auf Granulozyten zu finden ist, werden von den für Menschen gefährlichen Krankheitserregern der Spezies *Neisseria*, sowie von *Haemophilus influenzae* und *Moraxella catarrhalis* spezifisch erkannt.

In dieser Arbeit wurde das spiralförmige, Gram-negative Bakterium *Helicobacter pylori* auf Bindung an die CEACAM-Rezeptorfamilie untersucht. *H. pylori* ist ebenfalls ein für Menschen gefährlicher Krankheitserreger, der sich im menschlichen Magenepithel ansiedeln kann. Es konnte gezeigt werden, dass das Bakterium an die Rezeptoren CEACAM1 und CEACAM5 bindet, nicht aber an die anderen beschriebenen Familienmitglieder CEACAM3, 4, 6, 7 und 8. Daneben wurden auch Bakterien auf Bindung untersucht, die zur Gattung *Helicobacter* zählen, jedoch zu einer anderen Art als *pylori* gehören. Diese Bakterien konnten nicht mit CEACAMs interagieren. Die Bindung scheint sich zudem auf die im Menschen vorkommenden CEACAMs zu beschränken, da keine CEACAM-orthologen Strukturen in anderen Arten als dem Menschen von *H. pylori* erkannt werden konnten. Dies legt nahe, dass die Fähigkeit an CEACAMs zu binden eine Anpassung der Spezies *H. pylori* an ihren menschlichen Wirt darstellt.

In dem Versuch, die bakteriellen Adhäsine zu identifizieren, die an CEACAMs binden, wurden Deletionsmutanten hergestellt. Damit konnte gezeigt werden, dass ausschließlich das äußere Membranprotein HopQ für die Interaktion mit CEACAM1 und CEACAM5 verantwortlich ist. Von HopQ existieren zwei Genvarianten, die in mehr (HopQI) oder weniger (HopQII) virulenten *H. pylori*-Stämmen vorkommen. Beide HopQ-Genvarianten waren in der Lage, mit CEACAMs zu interagieren, aber die CEACAM-Bindung war für HopQI deutlich stärker als für HopQII.

Die *H. pylori*-Stämme PMSS1 und SS1 unterscheiden sich lediglich dadurch, dass der Stamm SS1 durch die Maus passagiert wurde. Als die Stämme auf ihre CEACAM-Bindefähigkeit untersucht wurden, zeigte sich, dass der Tier-passagierte Stamm SS1 seine CEACAM-Bindefähigkeit weitgehend verloren hatte. Der Grund hierfür war, dass die Genvariante HopQI nicht mehr produziert wurde, während in dem Stamm PMSS1 sowohl ein funktionales HopQI als auch ein HopQII gebildet wurde. Dies deutet darauf hin, dass *H. pylori*-Stämme ihre CEACAM-Bindefähigkeit durch die Mauspassage verlieren, vermutlich weil die Bakterien CEACAM-orthologe Strukturen nicht binden können.

Eukaryotische Zelllinien wurden auf die Synthese von CEACAMs untersucht und es zeigte sich, dass in CEACAM-positiven Zelllinien die epithelialen CEACAMs 1, 5 und 6 immer zusammen produziert werden. Die Adhäsion spielt eine wichtige Rolle bei der Translokation des Effektorproteins CagA in die Wirtszellen durch *H. pylori*, um zelluläre Prozesse zu beeinflussen. *In-vitro*-Infektionen von CEACAM- transfizierten HEK293-Zellen mit Wildtypbakterien und einer Adhäsionsdefekten Bakterienmutante zeigten, dass die CagA-Translokationsfähigkeit vollständig von der Gegenwart des Adhäsins HopQ abhängig ist.

Kurz zusammengefasst konnte das Adhäsins HopQ als neuartiger bakterieller Virulenzfaktor identifiziert werden, der mit einigen zellulären Rezeptoren der CEACAM-Familie interagiert. Die in dieser Studie gewonnenen Ergebnisse zeigen interessante neue Einblicke in die Pathogenese des Bakteriums *H. pylori* und in das Zusammenspiel dieses pathogenen Bakteriums mit seinem menschlichen Wirt.

# Summary

---

Being involved in main cellular processes like cell-cell adhesion, homoeostasis and regulation of the immune system, the surface exposed receptor family of CEACAMs has been discovered to be a favorable target for viruses and many bacteria and used as entry site into the host. The epithelial members CEACAM1, CEACAM5 and CEACAM6 together with the granulocyte receptor CEACAM3 are recognized by human-specific pathogens like *Neisseria* species, *Haemophilus influenzae* and *Moraxella catarrhalis*.

In this work, *Helicobacter pylori*, the spiral shaped Gram-negative and pathogenic bacterium residing in the gastric epithelium of the human stomach was examined for CEACAM binding capability. *H. pylori* was found to interact with the CEACAM family members CEACAM1 and CEACAM5. The bacterium did not show adhesion to CEACAM3, 4, 6, 7 and 8. Moreover, it was discovered that non-*pylori Helicobacter* species were not able to adhere to the CEACAM receptors and that the recognition is species-specific, since *H. pylori* was unable to adhere to CEACAMs found in other species than humans. This suggests that the CEACAM binding capability is a specific adaptation of the species *H. pylori* to its human host.

In the attempt to identify the bacterial adhesin adhering to CEACAMs, the Hop family was targeted for deletion and HopQ was identified as the sole adhesin targeting CEACAM1 and CEACAM5. HopQ exists in two allelic forms which are produced in more (HopQI) or less (HopQII) virulent *H. pylori* strains. In this work, it was found that both allelic forms of HopQ were generally recognizing CEACAMs, but the CEACAM binding capability was markedly enhanced for the HopQI variant.

The two *H. pylori* strains PMSS1 and SS1 differed only by animal passage of SS1. It was discovered that the strain SS1 had lost its binding ability to a large extent. The sequencing results showed a downregulation of the *hopQI* genes in the animal-passaged strain, whereas the non-mouse-passaged strain was still producing both a functional HopQI and HopQII. This suggests that *H. pylori* strains may lose their CEACAM-binding capability by mouse passage, likely due to a possible failure in recognizing CEACAM orthologues present in animals.

When eukaryotic cell lines were examined for CEACAM production, in CEACAM-positive cell lines, the epithelial CEACAMs 1, 5 and 6 were always co-produced. One major role of the *H. pylori* adhesion to the gastric epithelium is the translocation of the effector protein CagA into the host cells to modulate cellular processes. Interestingly, the CEACAM production on cell lines was supporting the CagA translocation capability. *In vitro* infections of CEACAM transfected HEK293 cells with wildtype and adhesion-defective mutant *H. pylori* strains showed a complete dependency of the CagA translocation on the presence of the adhesin HopQ.

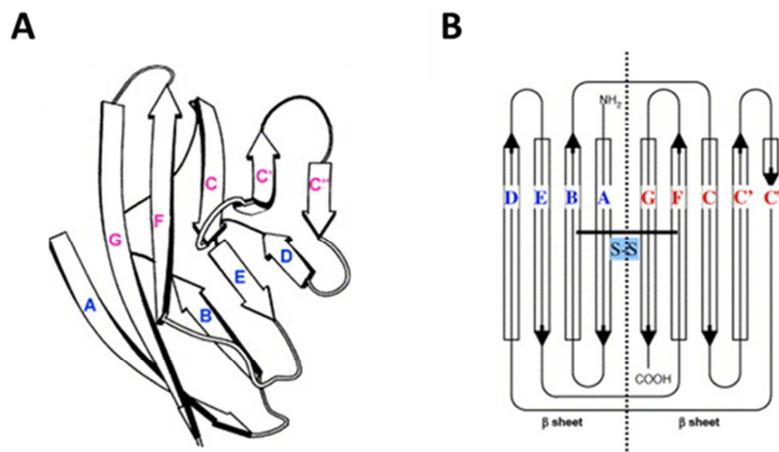
Taken together, the results obtained from this study identified the adhesin HopQ as a novel bacterial virulence factor interacting with the cellular CEACAM receptor family which might bring interesting new insights into *H. pylori* pathogenesis and the interplay between the pathogenic bacterium and its human host.

# Introduction

## 1 Immunoglobulin superfamily

### 1.1 COMPOSITION AND STRUCTURE

The Immunoglobulin superfamily (IgSF) has been one of the first superfamilies to be discovered and now represents one of the largest in vertebrate genomes. The IgSF members characteristically share the structural unit of Immunoglobulin-like (Ig-like) domains which are named due to their strong resemblance to antibodies (immunoglobulins), mostly in sequence, but also in function <sup>1</sup>. The domain consists of about 110 amino acids (aa) and forms a sandwich of two antiparallel  $\beta$ -sheets linked by disulfide bonds between cysteine residues of the B and F strands (Figure 1 A).



**Figure 1: Typical Ig domain of two  $\beta$ -sheets.**

**(A)**  $\beta$ -strands and loops that form the two antiparallel  $\beta$ -sheets. The strands were numbered from A to G. One  $\beta$ -sheet consists of strands A, B, E and D, while the other contains strands C, F, G and possibly C' and C''.

**(B)** The topology of the Ig domain is depicted. SS represents the conserved disulfide bonds between cysteine residues of the B and F strands (Barclay, 2003).

Ig domains are divided into four main types regarding their function and size: v, c, h and s type <sup>2</sup>. The v type (v for variable) or V-set is considerably similar to the variable, antigen binding region of antibodies. The v type comprises all strands of the two  $\beta$ -sheets including the C' and C'' strands (Figure 1 B) and presents the longest of Ig-like domains. The c type (c for constant) or C1-set is closely related to the constant region of antibodies which mediates effector functions. It lacks the C' and C'' strands in contrast to the v type <sup>3</sup>. The h type (h for hybrid) or C2-set is in

sequence pattern more similar to the v type, although it has a shortened domain length comparable to the c-type by lacking strand C''<sup>4</sup>. Besides these three types, the s type (s for switched) or I-set (intermediate) is devoid of strand D or C'' and typical for IgSF members that belong to the group of cell adhesion molecules (CAM) and surface receptors<sup>5</sup>.

## 1.2 FUNCTION

Members of the IgSF are found in a wide variety of proteins that differ largely in their function. Their ability to build rod-like structures and to interact specifically with other proteins makes them excellent candidates for CAMs or surface receptors<sup>6</sup>. IgSF domains are found in the nervous system (e.g. the neural cell adhesion molecule NCAM)<sup>7</sup>, but also in immune cells or cells that are involved in reproduction<sup>1</sup>. They are typically occurring in antigen receptors (T-cell receptor and antibodies<sup>8 9</sup>), in antigen presenting molecules (e.g. MHC class I and II)<sup>10</sup>, in coreceptors and costimulatory molecules of the immune system (e.g. CD28)<sup>11</sup> and even in muscle cells (e.g. Titin)<sup>12</sup>.

Proteins belonging to the IgSF are mostly multi-domain structures in which the domains are arranged like beads on a string. They can be divided into three groups regarding their Ig domain percentage: The first group containing members of the CEA gene family consists exclusively of Ig domains. The second group of IgSF proteins is characterized by one to four Ig domains and additional extracellular domains that show similarities to fibronectin type III (FNIII domains), such as NCAM. In a third group proteins possess other domains than FNIII domains besides Ig domains<sup>13</sup>.

## 2 Carcinoembryonic antigen gene family

One subgroup of the IgSF is represented by the carcinoembryonic antigen gene family (CGF) that is further divided into the carcinoembryonic antigen related cell adhesion molecules (CEACAMs) and the pregnancy-specific glycoproteins (PSG) <sup>14</sup>.

The human CGF is organized on chromosome 19 in contiguous clusters and consists of 35 genes. Only 21 of these genes are coding for proteins <sup>15</sup>. All family members are well conserved in their structure. One N-terminal IgV-like domain (V-set) is followed by several IgC-like domains (C2-set) on the extracellular part of the protein. The IgC-like domains can again be distinguished in subset A and B that are alternating in the organization of the CGF members <sup>16</sup>. The IgV-like domain is devoid of disulfide bonds between cysteine residues in the B and F strands which makes the domain more flexible and facilitates the adhesion to the domain <sup>17</sup>.

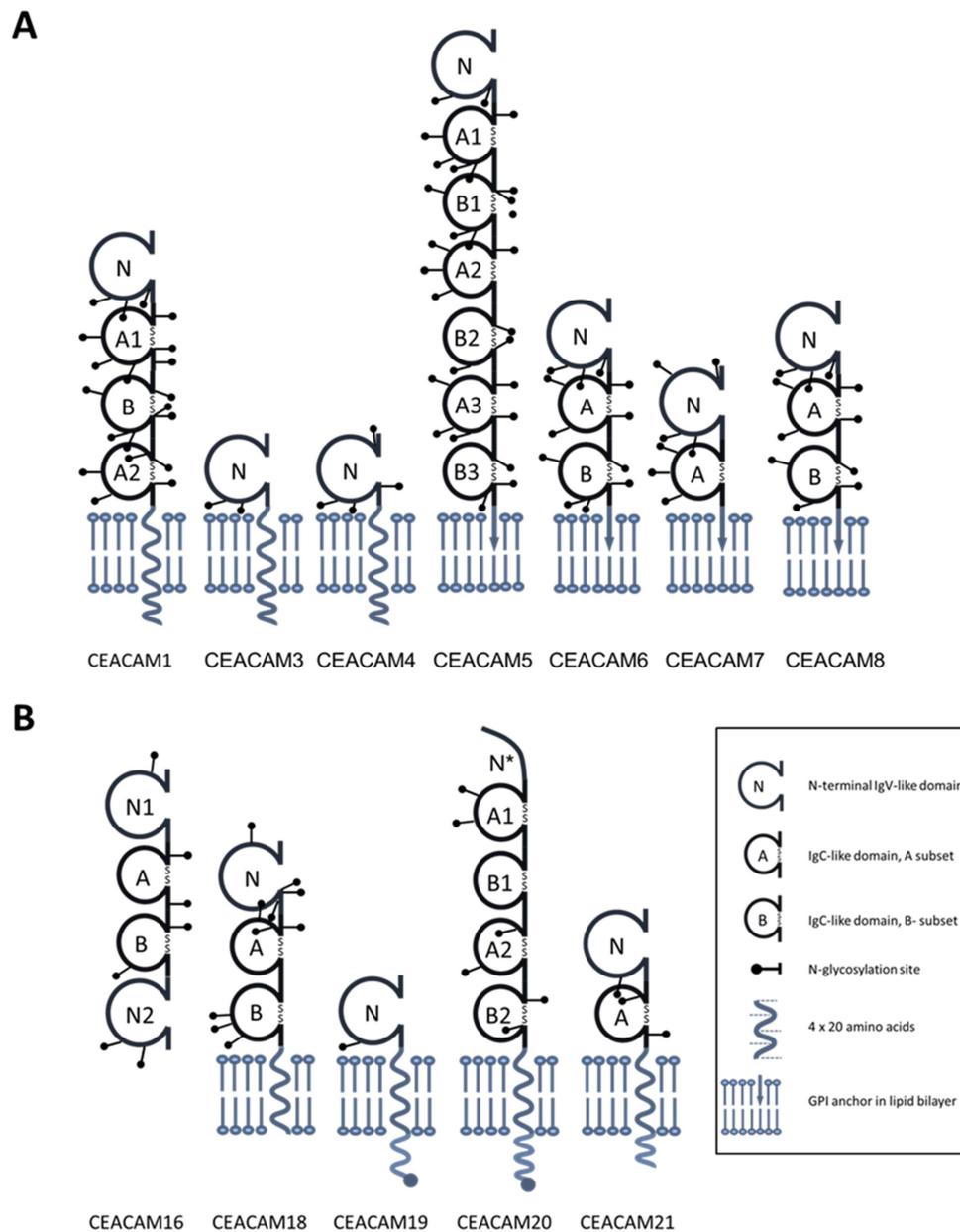
The family members are highly glycosylated proteins (up to 60 – 70% by weight). Glycosylation occurs as posttranslational modification <sup>18</sup> being beneficial for the interaction with binding partners by reduction of movement as well as for the binding specificity by protection of the binding site against non-specific binding <sup>19</sup>.

Members of the CGF are involved in various processes as cell-cell adhesion, pregnancy, immunity, neovascularization, regulation of insulin homeostasis and carcinogenesis <sup>20</sup>.

### 2.1 CEACAM FAMILY

The CEACAM subgroup of the CGF comprises 12 proteins. The well-studied members are CEACAM1, CEACAM3, CEACAM4, CEACAM5, CEACAM6, CEACAM7 and CEACAM8 (Figure 2 A). The family members from CEACAM16 to CEACAM20 have been identified more recently and not been characterized so well (Figure 2 B).

The number of IgC-like domains varies largely in the CEACAM group. Zero (e.g. CEACAM3) to six (e.g. CEACAM5) constant IgC-like domains are found on the extracellular part of CEACAM members. CEACAM5, CEACAM6, CEACAM7 and CEACAM8 are linked to the membrane by a glycosylphosphatidylinositol- (GPI-) anchor <sup>21</sup>, whereas the other members are type I membrane proteins containing a single transmembrane domain. The cytoplasmic domain of the transmembrane passing CEACAMs can contain either an immunoreceptor tyrosine-based activation motif (ITAM) or immunoreceptor tyrosine-based inhibitory motif (ITIM).



**Figure 2: CEACAM family members.**

The N-termed loops indicate IgV-like domains, the other spheres represent IgC-like domains which are stabilized by disulfide bonds (S-S). The spirals show the transmembrane domains which are followed by a cytoplasmic region. GPI-anchored proteins are depicted in the form of an arrow ending in the lipid bilayer. (A) CEACAM family members from CEACAM1 to CEACAM8 are depicted. The extracellular part of CEACAM3 and 4 consists of a single IgV-like domain. (B) CEACAM family members from CEACAM16 to CEACAM21 are depicted. CEACAM16 is not membrane bound, but secreted. CEACAM19 consists of a single IgV-like domain. CEACAM20 only encodes for a partial IgV-like domain (N\*).

(modified from <http://www.carcinoembryonic-antigen.de>)

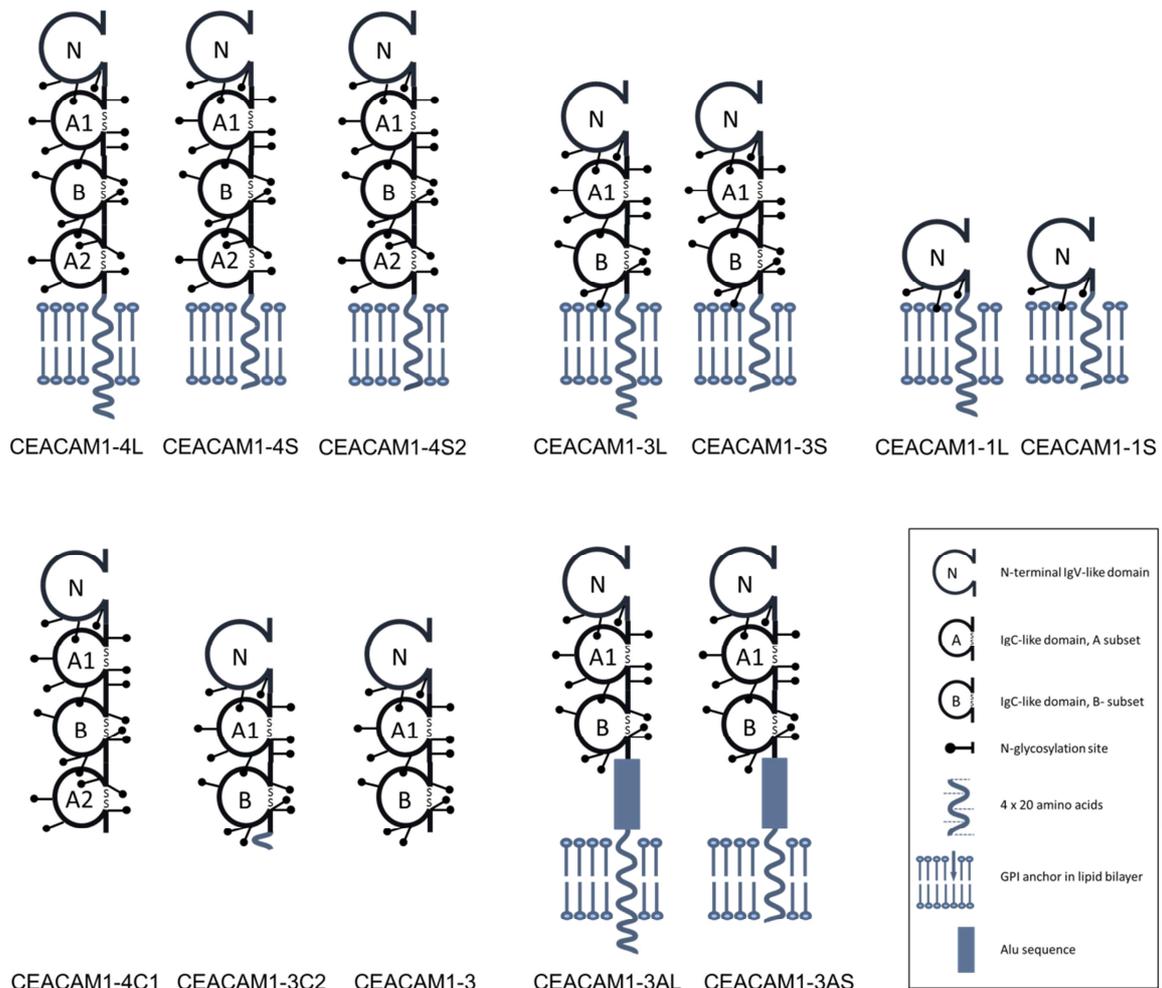
During the last decades, multiple nomenclatures have been introduced for CEACAMs. For example, clearly characterized monoclonal antibodies (mAb) of the CD66 cluster recognize diverse members of the CEACAM family and the antibodies' designations were used as a nomenclature for the corresponding CEACAM members. Thus, CEACAM1, CEACAM8, CEACAM6, CEACAM3 and CEACAM5 are also known as CD66a, CD66b, CD66c, CD66d and CD66e, respectively <sup>22</sup>. In 1999, a standardized nomenclature was introduced to facilitate the research on CEACAMs <sup>23</sup>.

The cell distribution of CEACAMs is diverse. Whereas CEACAM1 is produced on virtually every cell type like epithelial, endothelial, lymphoid and myeloid cells, the other members are more cell-specific. CEACAM3, 4, 6 and 8 are predominantly found on granulocytes. CEACAM5, 6 and 7 are the main CEACAMs present on epithelial cells. Little is known about the cell distribution of CEACAM16 to CEACAM21 <sup>24</sup>.

Various functions have been described for CEACAM members. They are involved in main cellular processes like homeostasis, intercellular adhesion and the regulation of immune cells. Given the central role of CEACAMs in cells, pathogenic bacteria have been discovered to exploit CEACAM1, CEACAM3, CEACAM5 and CEACAM6 as entry site into the host <sup>25</sup>. The structural and functional characteristics of these CEACAMs are discussed in more detail in the following section.

### CEACAM1

CEACAM1, formerly known as biliary glycoprotein (BGP), C-CAM and CD66a, was first identified in adult rat hepatocytes <sup>26</sup>. A total of 12 splice variants of CEACAM1 have been identified so far in humans (Figure 3). The CEACAM1 isoforms differ both in the number of extracellular IgC-like domains and in their binding activity. The number after CEACAM1 indicates the number of Ig-like domains, whereas the letter that follows indicates the presence of either a long (L, 75 aa) or a short (S, 14 aa) cytoplasmic tail generated by alternative splicing of exon 7 <sup>27</sup>. The L-isoform contains two ITIM motifs and generally transmits inhibitory signals in cells. Isoforms with a short cytoplasmic tail lack both ITIM sequences, but, like the L-isoform, have binding sites for effector proteins such as calmodulin, tropomyosin and globular actin, indicating an interaction with the cytoskeleton <sup>28 29</sup>. Furthermore, CEACAM1 affects T-cell cytokine production, proliferation and cytotoxic activity in an inhibitory way <sup>30</sup>. CEACAM1 is involved in angiogenesis <sup>31</sup>, in the regulation of B-cells <sup>32</sup> and natural killer (NK) cell function <sup>33</sup>. CEACAM1 is also important for the maturation of dendritic cells (DC) <sup>34</sup>.



**Figure 3: CEACAM1 isoforms by alternative splicing.**

The N-termed loops indicate IgV-like domains, the A- and B-termed spheres represents IgC-like domains which are stabilized by disulfide bonds (S-S). Except for the soluble CEACAM1-4C1, CEACAM1-3C2 and CEACAM1-3, the CEACAM1 isoforms are membrane bound. The number after CEACAM1 indicates the number of extracellular Ig-like domains, whereas the letter that follows this indicates the presence of either a long (L) or a short (S) cytoplasmic tail. (modified from <http://www.carcinoembryonic-antigen.de>)

In most cases both the L- and S-isoform are present on cells <sup>35</sup>, but the ratio between the isoforms differ. In mouse B-cells, it has been shown that the isoforms are produced with approximately equal ratios, whereas in mouse T-cells, the S-isoform presents only a minor part compared to the L-isoform <sup>27</sup>. This strongly outbalanced ratio between the two isoforms might influence the T-cell signaling outcome suggesting an ability of the coproduced S-isoform to fine tune the inhibitory properties of CEACAM1-L on T-cells <sup>36</sup>.

The formation of homodimers and heterodimers is essential for the functioning of CEACAM1. It can either form oligomers by *trans* (antiparallel) or *cis* (parallel) interactions. Homodimeric

*trans*-oligomerization is achieved by interaction of CEACAM1 IgV-like domains between cells which regulates cell-cell adhesion <sup>37</sup>. CEACAM1 can also build heterodimeric *trans*-interactions with selectins and other CEACAM family members, e.g. CEACAM5 and CEACAM8, respectively <sup>38</sup>. *Cis* dimerization of CEACAM1 leads to changes in the cells signaling outcome. Heterodimeric *cis*-interactions have been recently reported for CEACAM1 and TIM-3 (T-cell immunoglobulin domain and mucin domain-3). TIM-3 is an activation-induced inhibitory protein on T-cells that is involved in tolerance and induces T-cell exhaustion in chronic viral infection and cancer <sup>39</sup>. Furthermore, CEACAM1 interacts with  $\beta$ 3 integrin. This depends on phosphorylation of Tyr488 in the ITIM motif of the CEACAM1 cytoplasmic domain <sup>40</sup>.

Homodimeric *cis* dimerization leads to the formation of clusters which allow the recruitment of Src family kinases and the SRC homology 2-domain (SH2) containing protein tyrosine phosphatase 1 and 2 (SHP1 and SHP2). By the phosphorylation of both ITIM motifs, CEACAM1 is accessible for the SH2 domains of SHP1 as docking sites <sup>27</sup>.

Several factors have influence on the state of CEACAM1 and can shift the balance between the monomeric and dimeric form. An elevated  $Ca^{2+}$ -level leads to the formation of  $Ca^{2+}$ -calmodulin complexes and the subsequent dissociation of dimeric CEACAM1 *in vitro* <sup>41</sup>. In contrast, the transmembrane domain of CEACAM1 promotes the transition towards oligomerization of CEACAM1 with its (432)GxxxG(436) motif supporting helix-helix interactions. This basal state is overcome by calmodulin binding to CEACAM1, since the binding overlaps other possible binding sites <sup>42</sup>. The IgC-like domains may also contribute to binding and/or signaling events of CEACAM1, since the domains stabilize the dimerization of CEACAM1 <sup>43</sup>.

### CEACAM3

CEACAM3 is known as CD66d and CEA gene family member 1 (CGM1). Its single IgV-like domain is closely related to CEACAM1. The 76 aa long cytoplasmic tail contains an ITAM motif which, in contrast to CEACAM1, transmits activation signals <sup>44</sup>. CEACAM3 is exclusively present on human granulocytes and takes part in the phagocytosis of human-specific pathogens by the innate immune system. CEACAM3-bound bacteria are efficiently recognized in an opsonin-independent manner and internalized. CEACAM3 stimulates the small GTPase Rac <sup>45</sup> and activates protein tyrosine kinases (PTKs) of the Src family upon bacterial engagement <sup>46</sup>. Together with its downstream effector, the WAVE complex, Rac promotes the formation of actin-based lamellipodia which leads to a rapid internalization of CEACAM3-bound *Neisseria* <sup>47 48</sup>. By the binding of phosphatidylinositol 3'-kinase (PI3K) to the cytoplasmic domain of CEACAM3 on *Neisseria*-infected phagocytes the production of reactive oxygen species is stimulated leading to the subsequent intracellular degradation of CEACAM-bound bacteria <sup>49</sup>. No endogenous ligand has yet been identified and no homologue has been found in other mammalian species.

CEACAM3 is regarded as an evolutionarily new and optimized receptor to operate against the exploitation of epithelial CEACAMs by bacteria <sup>50</sup>.

### CEACAM5

CEACAM5 or CEA has formerly been described as CD66e and was identified as tumor associated antigen in human colon cancer <sup>51</sup>. CEACAM5 contains the highest number of extracellular IgC-like domains and is attached to the membrane by a GPI-anchor <sup>50</sup>. The CEACAM members comprising a semi-penetrating anchorage to the membrane have only been discovered in the group of primates, and an orthologue of CEACAM5 exists in the rhesus monkey <sup>52</sup>. CEACAM5 is exclusively produced on epithelial cells in the gastrointestinal tract, the nasopharynx, the lung, the urogenital tract and sweat glands and takes part in signal transduction processes <sup>53</sup>. Given the fact that CEACAM5 is GPI-anchored, it has not completely been resolved how it exhibits its signaling function without a membrane spanning and cytoplasmic domain. There is first evidence that CEACAM5 needs a distinct lipid composition to be capable of signaling events <sup>54</sup>. The GPI-anchor leads to the constitutive localization in detergent-resistant membrane microdomains which function as platforms for protein trafficking and intracellular signaling <sup>55</sup>. Without stimulation, CEACAM5 exists as a monomer or as nanoclusters of 5 – 10 nm diameters. After stimulation, clusters of 40 – 300 nm might form dependent on the composition of sphingolipids, cholesterol and actin. This leads to temporary immobilizations in the clusters and subsequent activation of signaling cascades by recruited G $\alpha$  proteins, Src family kinases and PLC $\gamma$  <sup>56</sup>. CEACAM5 also functions as mediator of metastasis and inhibits anoikis, a specific type of apoptosis triggered by the absence of ECM-cell contacts. By binding of heterogeneous nuclear ribonucleoprotein M (hnRNP M) on liver Kupffer cells, CEACAM5 leads to the activation and production of pro- and anti-inflammatory cytokines including IL-1, IL-10, IL-6 and TNF-alpha <sup>57</sup>. In addition to that, CEACAM5 inhibits NK cell cytotoxicity in a MHC class I independent way by heterodimeric interactions with CEACAM1 on NK cells <sup>58</sup>. By this, tumor cells are protected against the recognition and elimination by the immune system <sup>55</sup>.

### CEACAM6

CEACAM6 is known as CD66c and non-specific cross reacting antigen (NCA) and, like CEACAM5, is attached to the membrane by a GPI-anchor. It belongs to the epithelial CEACAM members that are recognized by pathogens, but is also found on human granulocytes as well as on B- and T-lymphocytes. It is involved in many crucial cellular events such as migration, invasion and tumorigenicity <sup>59</sup>.

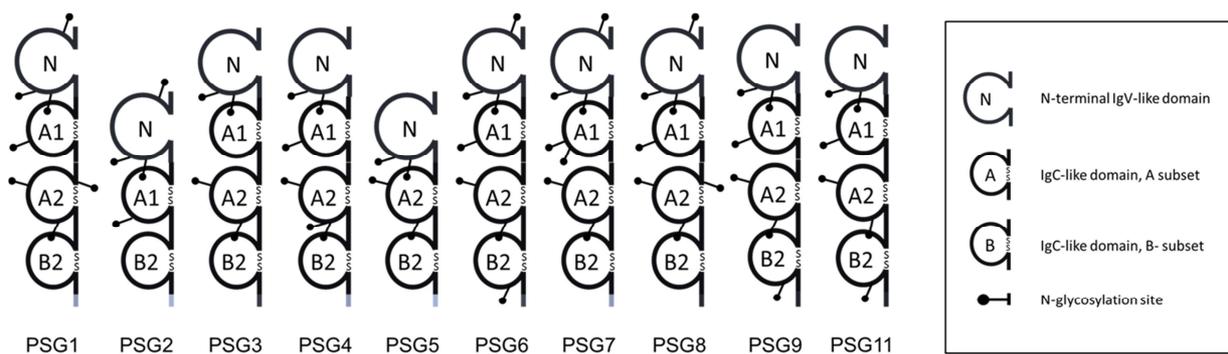
### Other CEACAMs

Other CEACAM members have not been found to interact with pathogens. CEACAM4 (or CGM7) has originally been identified from human peripheral leukocytes as a CEACAM6-related molecule on human granulocytes<sup>60</sup>. Recently, CEACAM4 has been identified as an orphan receptor regarding the lack of identified ligands for the IgV-like domain. Its cytoplasmic domain contributes to phagocytosis and elimination of pathogens<sup>61</sup>. CEACAM7 (or CGM2) is attached to the membrane by a GPI-anchor and typically found on epithelial cells. CEACAM8, also known as CD66b, is produced on human granulocytes as well as on T- and B-lymphocytes and linked to the membrane by a GPI-anchor. Recent studies reported the release of soluble CEACAM8 as a mechanism to inhibit the toll-like receptor 2 (TLR2)-triggered immune response<sup>62</sup>. The secreted CEACAM16 is the only CEACAM member with two IgV-like domains flanking two IgC-like domains and is associated with hearing in the inner ear<sup>63</sup>. The IgV-like domain of CEACAM18 is followed by two IgC-like domains. CEACAM19 has been found to be overexpressed in breast cancer tissue<sup>64</sup>. CEACAM20 lacks a complete IgV-like domain and contains an ITAM motif. CEACAM20 was found to be co-expressed with CEACAM1 in adult prostate tissue<sup>65</sup>. CEACAM21 consists of one IgV-like domain and one IgC-like domain that are followed by a transmembrane domain and cytoplasmic tail.

## 2.2 PREGNANCY-SPECIFIC GLYCOPROTEINS

PSGs are expressed during pregnancy and exclusively found in the placental trophoblasts<sup>14</sup>. A group of 10 members named PSG1-9 and PSG11 has been identified for humans (Figure 4). PSGs have been also discovered in non-human primates, rodents and bats, although the orthologous proteins differ strongly in their structural composition. Human family members are homogeneous consisting of one IgV-like domain and two to three IgC-like domains. The family members are secreted due to their carboxy-terminal tail<sup>15</sup>.

The synthesis of single PSG family members varies throughout pregnancy, but especially during late pregnancy, they are the most abundant fetal proteins found in maternal blood. Low levels of PSGs in blood have been related to complications in pregnancy. PSGs are assumed to protect the semi-allogeneic fetus against the maternal immune system and to access maternal resources during pregnancy<sup>66 67</sup>.



**Figure 4: PSGs.**

**The human subgroup of the CGF is very homogenic in its composition. One IgV-like domain is followed by two to three IgC-like domains that can be further distinguished in an A and B subset. All PSGs are secreted and only found during pregnancy in the placental trophoblast. (modified from <http://www.carcinoembryonic-antigen.de>)**

Most PSG members share the common tripeptide sequence Arg-Gly-Asp (RGD) in the N domain. The RGD motif is recognized by integrins which led to the hypothesis that integrins interact with PSG family members. Similar to the effect of snake venom disintegrins, which block the association of integrins with other partners, a binding between integrins and PSGs may lead to the disruption of cell - ECM interactions and the blocking of integrin-mediated cellular functions

68.

PSG1, the highest expressed PSG family member, is also known as CD66f. Recent studies identified for PSG1 a KGD motif instead of the RGD motif. The KGD motif is also found in the snake venom disintegrin barbourin which specifically inhibits fibrinogen binding to the  $\alpha$ IIb $\beta$ 3 integrin on platelets. Studies showed that, analogous to barbourin, PSG1 binds  $\alpha$ IIb $\beta$ 3 and antagonizes the platelet – fibrinogen interaction<sup>69</sup>. PSG1 is vital for the development of the fetus. Low levels have been associated with proinflammatory and anti-angiogenic phenotypes, like 'small-for-gestational age'-fetus. PSG1 induces the production of anti-inflammatory cytokines such as interleukin-10 (IL-10) and transforming growth factor beta-1 (TGF $\beta$ 1)<sup>70 71</sup>. TGF $\beta$ 2 is also assumed to be regulated by PSG1. TGF $\beta$ s regulate many biological processes essential for pregnancy success including trophoblast invasion and proliferation, angiogenesis, extracellular matrix formation and tolerance to the fetal semi-allograft<sup>66</sup>.

PSG1 was found to bind cell surface heparan sulfate (HS) and chondroitin sulfate (CS) proteoglycans, as well as the four members of the Syndecan family (Syndecan1-4). Proteoglycans consist of a protein core and covalently attached glycosaminoglycan (GAG) chains. During pregnancy, the growth and survival of the fetus depends on the access to increased placental blood flow. Syndecans are essential in angiogenesis and have been shown to regulate TGF $\beta$  activity. They are involved in suppressing the production of proinflammatory cytokines<sup>72</sup>. Through binding to Syndecans and other cell surface proteoglycans on endothelial cells, PSG1 induces the subsequent TGF $\beta$ -mediated induction of tube formation<sup>73</sup>.

Taken together, PSGs regulate diverse cellular processes like angiogenesis and modulate the maternal immune system which is essential for pregnancy success.

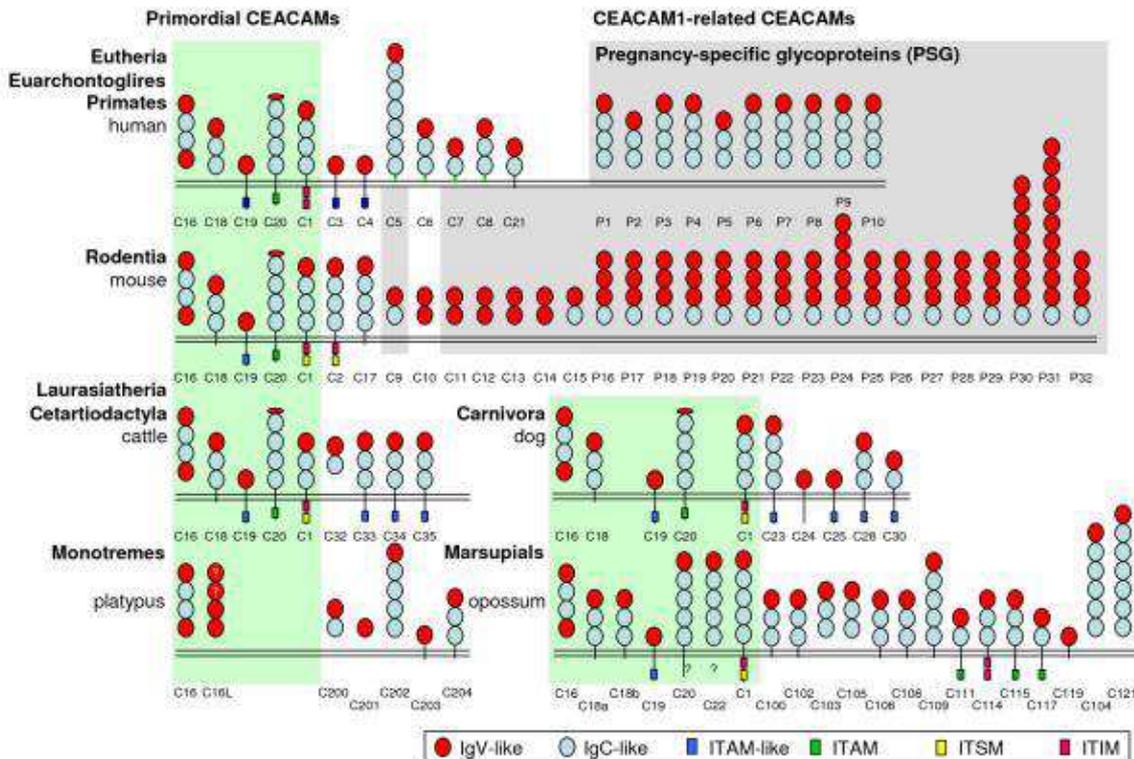
## 2.3 CEACAMS IN OTHER SPECIES AND EVOLUTION

The CGF has developed from one common ancestor gene. During time, gene duplications have led to a considerable diversification in protein structure, expression and function of CEACAMs and PSGs. Orthologues of CEACAMs have been identified in a variety of vertebrate species. CGF members are not conserved in their whole gene sequence, but the IgV-like domain is found to be highly conserved between species and is sufficient to identify orthologous structures <sup>74</sup>.

CEACAM1 and the recently identified CEACAM16, CEACAM18, CEACAM19 and CEACAM20 have been classified as primordial CEACAMs, since they have been discovered in all analyzed species. They might represent the basic set of CGF members before the occurrence of a rapid expansion in species-specific members. Monotremes show the only exception to this with merely a CEACAM16 orthologue identified <sup>75</sup>. Figure 5 shows the distribution of CGF members in selected mammals <sup>76</sup>.

Orthologues for GPI-anchored CEACAMs are exclusively found in primates. Studies could show that, by few mutational modifications, the GPI-anchor is generated of the transmembrane domain exon and, therefore, might have easily evolved in a CEACAM1-like ancestor. The development of CEACAMs equipped with a GPI-anchorage out of former membrane-bound members allows the secretion of the proteins in the extracellular milieu which might lead to a radical change or the expansion in molecular function of CEACAMs <sup>77</sup>.

So far, CEACAM3 and CEACAM4 have been identified merely in humans. The IgV-like domain of CEACAM3 is highly similar to the CEACAM1 domain (88% aa identity). Closely related molecules to the N-terminal domain of CEACAM1 like CEACAM3, CEACAM5 and CEACAM6 might act as granulocyte specific or decoy receptors for the uptake of invading pathogens and to regulate immune responses more precisely <sup>78</sup>.



**Figure 5: CGF members in mammals.**

The diagram shows the distribution of the CGF in mammals. CEACAMs shaded in light green represent the basic set of primordial members. Orthologues have been detected in all of the analyzed genomes, except for monotremes which show only an orthologous structure for CEACAM16. Due to rapid diversification, species-specific CEACAMs have evolved from a CEACAM1 ancestor protein (CEACAM1-related CEACAMs). PSGs are shaded in light grey and have only been found in primates and rodents. CEACAMs linked to the membrane by a GPI-anchor occur merely in primates. IgV-like domains (red) and IgC-like domains (light blue) are depicted by circles. (Kammerer and Zimmermann, 2010)

Most of the species possess only one *CEACAM1* gene, whereas species like mouse, cattle and rat carry two gene copies. In mouse, the translated CEACAM1 proteins named murine CEACAM1<sup>a</sup> and CEACAM1<sup>b</sup> differ in their IgV-like domain. The murine CEACAM1<sup>a</sup> isoform is bound by the mouse hepatitis virus with high affinity, whereas the CEACAM1<sup>b</sup> isoform is only recognized to a minor degree. The sequence identity between the IgV-like domains of human CEACAM1 and the murine CEACAM1<sup>a</sup> is low with an aa similarity of about 45%, whereas the CEACAM1<sup>b</sup> isoform shows a conservation of approximately 55% to the human variant<sup>79</sup>. The cytoplasmic tail of the murine CEACAM1 isoforms does not contain two ITIM motifs, but one ITIM and an immunoreceptor tyrosine-based switch motif (ITSM). The ITSM motif can send both stimulatory and inhibitory signals, while ITIM motifs transmit only inhibitory signals<sup>76</sup>.

The CEACAM1 isoforms found in cattle consist of one IgV-like and two IgC-like domains (A1, A2) and lack the B subtype of IgC-like domains found in humans and mouse. The sequence encoding for the B domain is missing in the bovine *CEACAM1* locus. Besides that, the bovine CEACAM1 has the typical structural composition described for the murine CEACAM1: It is membrane bound and has a cytoplasmic tail with an N-terminal ITIM and a C-terminal ITSM motif. The bovine IgV-like domain exhibits an aa sequence similarity of 54% to the human variant and 38% compared to the murine variant <sup>80</sup>.

The CEACAM1 orthologue found in the dog is composed like the human version (N, A1, B, A2). The cytoplasmic tail contains an N-terminal ITIM and a C-terminal ITSM motif like in mouse, rat and cattle. The canine IgV-like domain shows a similarity in aa sequence of 56% to the human CEACAM1 and 42% to the murine CEACAM1 isoform. The IgV-like domain of the canine CEACAM1 and CEACAM28 differ merely in two aa and, therefore, might interact with the same ligands and pathogens. The canine CEACAM28 is found on canine granulocytes and assumed to have evolutionary coevolved with CEACAM1 as inhibitory and activation receptors for the regulation of T-cell responses similar to the human CEACAM1/CEACAM3 receptors <sup>75</sup>.

## 2.4 CEACAM AND CANCER

In 1965, CEACAM5 was identified as a 180 kDa, tumor-associated antigen in colon carcinoma <sup>51</sup> and, since then, became an important marker for the recognition of colorectal and other tumors. In addition to that, the preoperative CEACAM5 serum level is analyzed as a measure for the estimated overall survival rate of patients with colon cancers and to stage cancer patients <sup>81</sup>. CEACAM5 protects tumorigenic cells from the immune system and promotes metastasis by inhibition of NK cell-mediated killing <sup>58</sup> and anoikis <sup>57</sup>.

Tumor-promoting effects have also been reported for other members of the CEACAM family. Especially well studied are CEACAM1 and CEACAM6. Besides CEACAM5, CEACAM6 is overexpressed in about 70% of all human tumors <sup>82</sup>, involved in the maintenance of the gastrointestinal tissue architecture and contributes, when overexpressed, to tumor formation in the colon by inhibition of differentiation and anoikis <sup>83</sup>.

Since mice do not have genes for CEACAM5 and CEACAM6, transgenic mouse models were established to analyze the effect of CEACAMs on tumor growth *in vivo*. Since transgenic CEACAM5 mice did not show an influence on tumorigenesis at all <sup>84</sup>, more promising results were obtained by another mouse model, the CEABAC model. A 187 kDa bacterial artificial chromosome (BAC) transgene was introduced into mice which codes for human CEACAM3, CEACAM5, CEACAM6 and CEACAM7 and, additionally, comprises the regulatory elements of the CEACAMs <sup>85</sup>. The CEABAC transgenic mice were treated with azoxymethane (AOM) in order to induce colon tumor formation and showed more than a two fold increase in mean tumor load relative to their wildtype littermates. The level of CEACAM5 and CEACAM6 increased by two- and 20-fold, respectively, and mimicked very closely the situation in human colon tumorigenesis <sup>86</sup>.

In contrast to CEACAM5 and CEACAM6, CEACAM1 is predominantly downregulated in early phases of cancers <sup>87</sup>. The increased synthesis of CEACAM1 in cancer cells was reported to reduce the tumorigenic phenotype both *in vitro* and *in vivo* in an ITIM-dependent manner <sup>88</sup> <sup>89</sup>. In aggressive and more progressed cancers, CEACAM1 is found to be overexpressed and high levels of the L-isoform were correlated with metastatic spread <sup>90</sup>.

Therapeutic strategies have been evaluated to target CEACAMs in cancer patients. Both, the siRNA mediated silencing and monoclonal antibodies (mABs) directed against CEACAM1, CEACAM5 and/or CEACAM6, are promising approaches for immunotherapies to increase the overall survival rate of cancer patients <sup>91</sup>.

## 2.5 CEACAM RECOGNITION BY PATHOGENS

Surface exposed structures like CEACAMs are target structures of many pathogenic viruses and bacteria. From all the CEACAM family members, the N-terminal IgV-like domain of the epithelial CEACAM1, CEACAM5 and CEACAM6 together with the granulocyte receptor CEACAM3 have been found to be recognized by human-specific pathogens (Figure 6).

Bacterial species	Comment	Adhesin type	CEACAM binding determinants	CEACAM1	CEACAM3	CEACAM5	CEACAM6	References
<i>Neisseria gonorrhoeae</i>	Pathogen: gonorrhoea, pelvic inflammatory disease, conjunctivitis	Opa (integral outer membrane protein)	Protein	Yes	Yes	Yes	Yes	Gray-Owen et al., 1997; Virji et al., 1996; Bos et al., 1997; Chen et al., 1997
<i>Neisseria meningitidis</i>	Pathogen: meningitis, septic shock	Opa (integral outer membrane protein)	Protein	Yes	Yes	Yes	Yes	Virji et al., 1996; Muenzner et al., 2000; de Vries et al., 1998
<i>Neisseria subflava</i>	Commensal: nasopharynx	Opa (integral outer membrane protein)	Protein	Yes	ND	ND	ND	Toleman et al., 2001
<i>Neisseria lactamica</i>	Commensal: nasopharynx	Opa (integral outer membrane protein)	Protein	Yes	ND	ND	ND	Toleman et al., 2001
<i>Moraxella catarrhalis</i>	Pathogen: otitis media	UspA1 (fibril like)	Protein	Yes	Yes	Yes	Yes	Hill et al., 2001; Hill et al., 2005
<i>Haemophilus influenzae</i> (encapsulated)	Pathogen: meningitis, septic shock	OmpP5 (integral outer membrane protein)	Protein	Yes	ND	ND	ND	Virji et al., 2000
<i>Haemophilus influenzae</i> (non-typable)	Pathogen: otitis media, pneumonia	OmpP5 (integral outer membrane protein)	Protein	Yes	ND	Yes	Yes	Virji et al., 2000; Klaile et al., 2013
<i>Escherichia coli</i>	Pathogen: enteric and urinary-tract infections	Fimbriae	Carbohydrate	ND	ND	Yes	Yes	Leusch et al., 1990; Leusch et al., 1991
Diffusely adhering <i>Escherichia coli</i> (DAEC)	Commensal: colon	Fimbriae	Carbohydrate	ND	ND	Yes	Yes	Leusch et al., 1990; Leusch et al., 1991
<i>Escherichia coli</i>	Pathogen: enteric disease	Afimbrial and fimbrial	ND	Yes	No	Yes	Yes	Berger et al., 2004
<i>Salmonella typhi</i>	Pathogen: typhoid fever	Fimbriae	Carbohydrate	ND	ND	Yes	Yes	Leusch et al., 1991
<i>Salmonella paratyphi</i> fever	Pathogen: paratyphoid fever	Fimbriae	Carbohydrate	ND	ND	Yes	ND	Leusch et al., 1991
<i>Salmonella java</i>	Pathogen: enteric fever	Fimbriae	Carbohydrate	ND	ND	Yes	ND	Leusch et al., 1991

**Figure 6: Bacterial species examined to bind CEACAM receptors.**

**Gram-negative bacteria have been characterized for their capability to interact with CEACAM family members. The corresponding adhesin is binding either to the protein or to the carbohydrate part of the CEACAMs.**

**ND (Not defined): The CEACAM-binding capability of the bacteria has not been analyzed;**

**Yes: Bacteria bind to the CEACAM family member; No: Bacteria have been examined for an interaction with the distinct CEACAM member, but do not show binding capability.**

**(modified from Gray-Owen et al., 2006)**

For pathogens, it is highly advantageous to target CEACAM1, since the receptor is produced on nearly every cell type and occurs in a variety of different species. This means that the evading organisms could easily access the host and also switch to new hosts <sup>76</sup>. Pathogens binding to CEACAM1 can influence both the innate and the adaptive immune response: *Neisserial* binding to CEACAM1 led to the inhibition of CD4<sup>+</sup> T-cells via ITIM signaling which further suppressed B-

cell responses<sup>92</sup>. In bronchial epithelial cells, it was shown that CEACAM1-binding provoked the downregulation of the innate immune system by suppression of TLR2 signaling<sup>93</sup>. In mice, CEACAM1 is the cellular receptor which is targeted by the mouse hepatitis virus (MHV).

Interestingly, the host has reacted to the exploitation of CEACAMs by the development of diverse counterstrategies against pathogens:

In humans, the N-terminal domain of CEACAM3 is nearly identical to CEACAM1, but mediates the opsonin-independent elimination of a restricted set of Gram-negative bacterial pathogens, including *Neisseria gonorrhoeae*, *Neisseria meningitidis*, *Haemophilus influenzae* and *Moraxella catarrhalis*. The cytoplasmic ITAM motif is essential in this process<sup>44</sup>.

The GPI-anchored CEACAM5 and CEACAM6 have also been suggested to be a defense mechanism of the host. By vesiculation of the microvilli formed by intestinal epithelial cells or hydrolysis of the GPI-anchor from the membrane, CEACAM-bound pathogens might be shedded from the cells to fend off pathogens<sup>94</sup>.

The IgV-like domain exon of CEACAM1 is the least conserved nucleotide sequence, when compared to different CEACAM members. Due to low conservation pressure, the N-terminal region of CEACAM1 is highly divergent between species and susceptible to genetic modifications which might have led to additional host strategies against invasion. The murine CEACAM1 aa sequence varies from other species in the first half of the CC' loop (aa 37 – 41) which was identified as part of the binding region for the mouse hepatitis virus. From the two distinct *CEACAM1* alleles in mice that have occurred due to gene duplication only CEACAM1<sup>a</sup> represents a high affinity receptor for the mouse hepatitis virus compared to CEACAM1<sup>b</sup><sup>80</sup>.

It is also proposed that soluble decoy receptors exist (for example CEACAM10 in rodents) which might allow pathogen binding due to the identical binding sequence, but which lead to no signaling outcome<sup>76</sup>.

## 3 *Helicobacter pylori*

### 3.1 DISCOVERY OF *H. PYLORI*

*H. pylori* is a Gram-negative, spiral shaped bacteria colonizing the gastric epithelium. It was postulated for a long time that microorganisms are not able to survive the acidic environment in the stomach. Despite this common opinion, Walter Krienitz and others had already observed spirochete bacteria in the human stomach of patients with gastric cancer <sup>95</sup>. However, the bacteria could neither be isolated from the tissue nor were they connected to any disease outcome. In 1984, Barry Marshall and John Robin Warren finally succeeded in culturing the bacteria which were named "*Helicobacter pylori*" due to their helical shape <sup>96 97</sup>. A few years later, Marshall and Warren were able to prove that *H. pylori* is associated with chronic gastritis and peptic ulcers <sup>98</sup>. In 2005, they received together the Nobel Prize in medicine for their achievement in the field of gastrointestinal pathogens.

### 3.2 *H. PYLORI* TRANSMISSION AND PREVALENCE

Usually, the acquisition of *H. pylori* occurs early in childhood and the transmission is assumed to be by close personal contact. Oral–oral and fecal–oral transmission are the most probable mechanisms for an infection with *H. pylori*, since the microaerophilic bacterium is not surviving outside the human body for long. When exposed to atmospheric oxygen and UV-light, the bacteria switch from actively dividing and swimming bacilli to the inactive form of cocci. Studies showed that *H. pylori* is not traceable in feces of healthy, but *H. pylori*-positive individuals. In contrast, the feces of *H. pylori*-positive individuals suffering from diarrhea were containing infectious bacilli as was the vomit of these people <sup>99</sup>.

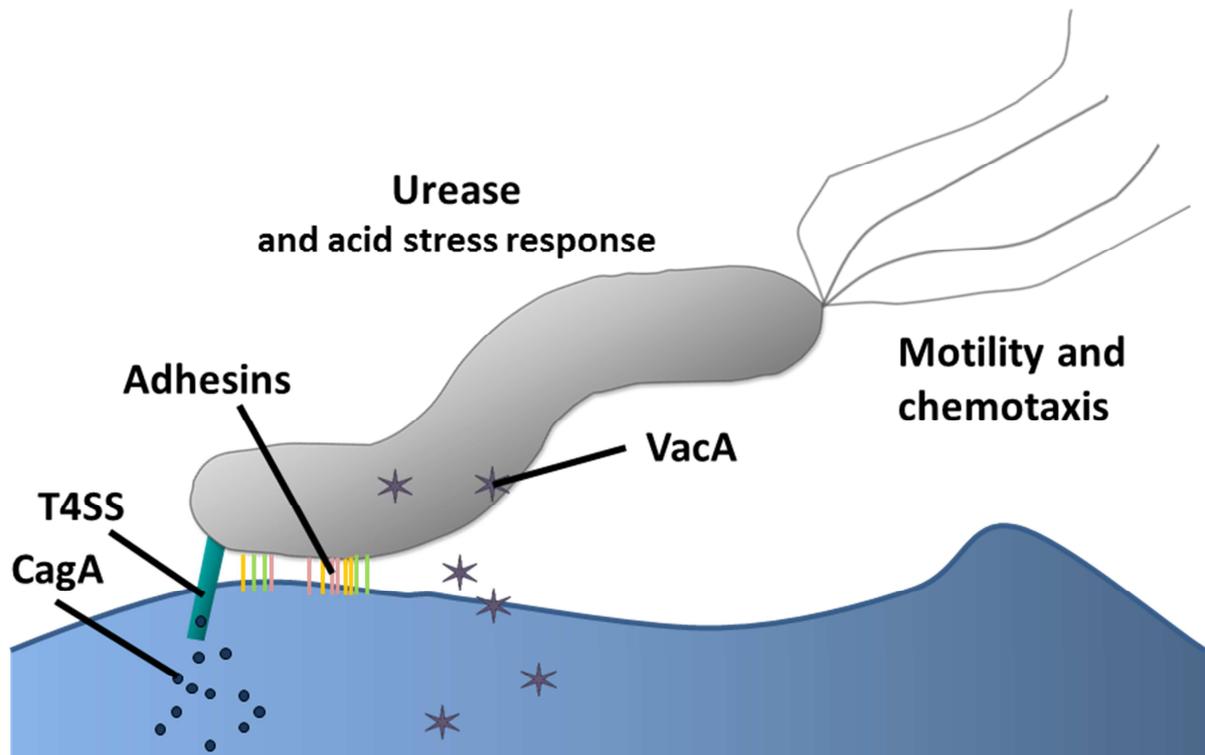
About 50% of the world's population is colonized by *H. pylori*, but only a small percentage of colonized individuals develop an *H. pylori*-related disease. The distribution of *H. pylori* between developed countries and developing countries varies strongly. Whilst in Western countries only 10% to 60% of the individuals are infected with *H. pylori*, in developing countries up to 100% of the population is positive for the bacteria. Risk factors for the bacterial occurrence are low socio-economic conditions like large family size <sup>100</sup>.

### 3.3 *H. PYLORI* AND DISEASE

*H. pylori* colonizes the antrum and the corpus of the gastric mucosa and its presence can be associated with severe pathologies. Three main phenotypes are caused by the bacteria. The most common phenotype is a mild pangastritis with little disturbance in gastric acid secretion which remains mostly asymptomatic. Secondly, the duodenal ulcer phenotype of about 15% of infected individuals has a more serious outcome. Patients show a severe antral predominant pattern of gastritis with elevated acid output and defective inhibitory control of gastric acid secretion which results in the development of peptic ulcers. The third phenotype that occurs in about 1% of infected individuals is characterized by a corpus predominant pattern of gastritis, gastric atrophy and hypo- and achlorhydria. Additionally, the risk for the development of gastric cancer (mucosa-associated lymphoid tissues (MALT) lymphoma and antral adenocarcinoma) is increased. The infection with *H. pylori* is the main risk factor for 92% of gastric cancers which causes 740 000 deaths per year. In 1994, the International Agency for Research on Cancer (World Health Organization) declared *H. pylori* as a type I carcinogen <sup>101</sup>. *H. pylori* eradication treatment comprises the combined intake of a proton pump inhibitor with the antibiotics amoxicillin and clarithromycin (French triple) or clarithromycin and metronidazole (Italian triple). When the therapy is not working, the quadruple therapy is used for the eradication, consisting of a proton pump inhibitor combined with tetracyclin, metronidazole and bismuth. The gastric cancer phenotype is especially prevalent in Asia <sup>99</sup>. The different disease manifestations are assumed to be caused by multiple factors like the virulence factors present in the *H. pylori* strains, host genetic characteristics, the environment as well as dietary factors <sup>102</sup>.

### 3.4 MAJOR VIRULENCE FACTORS

Many factors account for the virulence of *H. pylori* (Figure 7). The helical shape and the lophotrichous flagella help the bacteria to relocate from the acidic stomach lumen (pH~2.0) to the moderate mucosal surface (pH~7.0) through the viscous mucus layer <sup>103</sup>. Motility of *H. pylori* is an essential factor to survive in the acidic environment. Directed motility is achieved by chemotactic receptors which enable the bacteria to swim towards the more favorable gastric epithelium <sup>104</sup>.



**Figure 7: Major virulence factors of *H. pylori*.**

Directed motility by the lophotrichous flagella and chemotactic sensors help the bacteria to relocate from the acidic stomach lumen to the moderate mucosal surface. By the production of urease and acid stress responses, *H. pylori* is capable of neutralizing the acidic environment in the stomach for a short time. The cytotoxin CagA is injected into the host cells by the *cagPAI* type IV secretion system (T4SS) and leads to significant modifications regarding the cytoskeleton and signaling events in host cells. The exotoxin VacA is secreted in its active form to host cell membranes where it forms hexameric anion selective channels which cause the leakage of ions and nutrients. Adherence to the mucosal surface is critical to escape the mechanical clearance from the human stomach and to establish bacterial persistence.

(modified after Amieva and El-Omar, 2008)

Additionally, the bacteria produce large amounts of cytosolic and cell surface-associated urease. The enzyme is capable of neutralizing the acid for a short time by the breakdown of urea to ammonia and carbon dioxide so that the bacteria survive these conditions and are able to relocate to less acidic areas. Furthermore, several acid-stress responses enable as urease-independent mechanisms the transient toleration of mild acidic conditions <sup>99</sup>.

*H. pylori* strains can be distinguished into two main groups regarding their virulence: type I and type II. The more virulent type I strains are positive for the *cagPAI* pathogenicity island (*cagPAI*) and correlate with a s1 type of the vacuolating cytotoxin A (VacA) and a HopQ of the type I (HopQI), whereas the less pathogenic type II strains lack *cagPAI* and produce a s2 type of VacA and a HopQ of the type II (HopQII) <sup>105</sup>. The *cagPAI* encodes for the cytotoxin-associated gene A

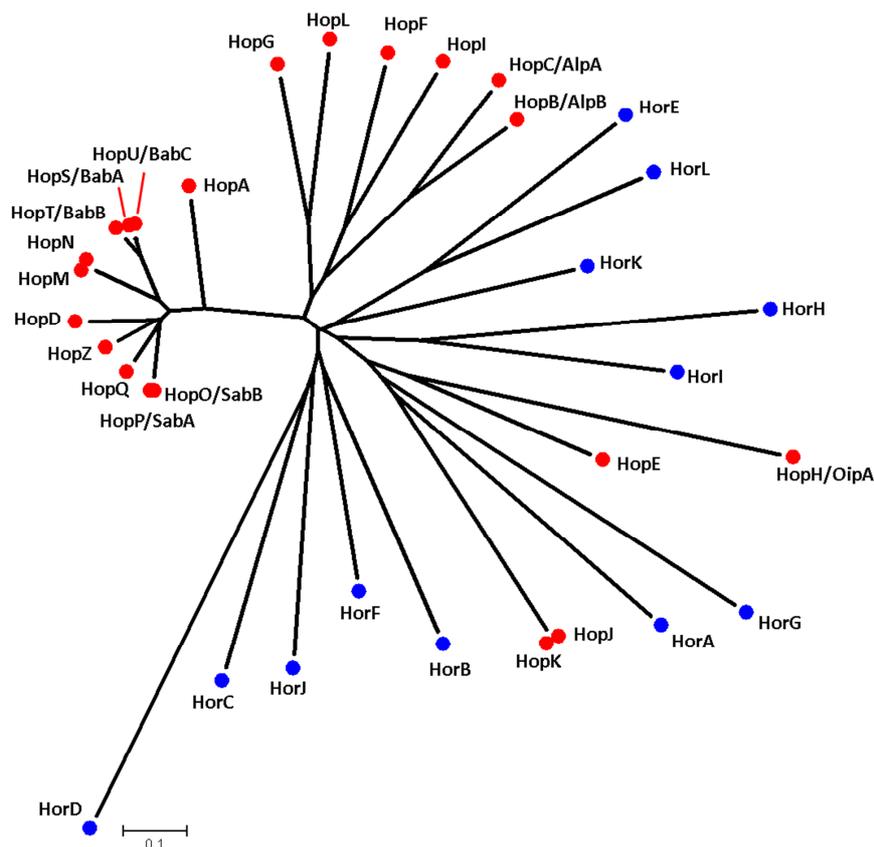
(CagA) and a type IV secretion system (T4SS) which injects the effector protein CagA into the cytoplasm of host cells. CagA is present in approximately 60% of Western clinical isolates and 95 – 100% of East Asian *H. pylori* strains<sup>106</sup>. The carboxy-terminal region of CagA varies among strains in its numbers of Glu-Pro-Ile-Tyr-Ala repeats (EPIYA)<sup>107</sup>. After translocation of CagA into host cells, the EPIYA motifs undergo tyrosine phosphorylation by cellular Src and c-Abl family kinases<sup>108</sup>. By this, intense cytoskeletal rearrangements are promoted which cause cell elongation (*hummingbird* phenotype) and increase cellular motility. Furthermore, the function of tight and adherens junctions is perturbed and signaling cascades of the host cells are influenced<sup>109</sup>. More EPIYA copies were associated with a higher virulence of the strain, since the strains show a stronger CagA tyrosine phosphorylation which results in a higher impact on the signaling and rearrangement events in host cells<sup>110</sup>. Additionally, CagA also interferes with host cells independent of phosphorylation. It was reported that the dimerization of non-phosphorylated CagA leads to the loss of cell polarity and cell shape<sup>111</sup>.

Besides CagA, the exotoxin VacA contributes to the virulence of *H. pylori* strains. It integrates in its active form into host cell membranes where it forms hexameric anion selective channels and causes the leakage of anions and other small molecules and subsequent osmotic swelling<sup>106</sup>. The cytotoxin is internalized into the host cell, targets mitochondria and leads to apoptosis of epithelial cells<sup>112</sup>. Furthermore, VacA inhibits T-cell activation and proliferation<sup>113</sup>. Unlike CagA, VacA is present in all *H. pylori* strains, but it shows allelic variations in the signal sequence of the amino terminal region (genotype s1 and s2) as well as in the middle region (genotype m1 and m2), which results in different virulence outcome. The more virulent VacA s1 genotype is associated with the secretion of active VacA and higher rates of ulcers and gastric cancer development compared to the s2 genotype. Most CagA-positive strains have a s1m1-genotype<sup>99</sup>. Increasing evidence suggests that there is a functional relation between CagA and VacA in the infection of the gastric epithelium by CagA interfering with VacA's cytotoxicity<sup>114 115</sup>.

The survival of the bacteria is also dependent on adherence, another important virulence factor. Due to the constant turnover in the stomach, the steady exocytosis of mucopolysaccharides and the secretion of gastric juices, the bacteria are always endangered to be cleared from the human stomach. Adherence to the mucosal surface is therefore critical for the persistence of the bacteria in the stomach and the colonization of the host<sup>116</sup>.

### 3.5 *H. PYLORI* ADHESINS

The adherence of *H. pylori* might serve multiple purposes. One major role of adhesion is assumed to be the delivery of the cytotoxin CagA into host cells. Besides this, it is also beneficial for bacteria to escape the mechanical clearance from the human stomach by the attachment to the gastric epithelium. In the long term, adherence also affects the establishment of bacterial persistence. The currently identified adhesins all belong to the group of the outer membrane proteins (OMP). The OMPs of *H. pylori* are divided into five paralogous families and all adhesins have been assigned to the first family containing the Hop (*Helicobacter pylori* outer membrane protein) (Figure 8, red) and Hor (*Helicobacter pylori* outer membrane protein related) proteins (Figure 8, blue) <sup>99</sup>.



**Figure 8: Hop and Hor family tree of OMPs.**

The OMPs can be distinguished in five families. The first family of OMPs is depicted comprising the Hop (red) and Hor proteins (blue). The putative adhesin branch of the Hop proteins (branch on the left) consists of 11 Hop proteins which share a strong sequence identity and contain the so far identified *H. pylori* adhesins like BabA, SabA, HopD and HopZ: (W. Fischer, unpublished)

The putative adhesin branch of the Hop family consists of 11 Hop proteins which share a strong sequence identity and most of the identified adhesins like BabA, SabA, HopD and HopZ are found in this group <sup>117</sup>. The expression level of the OMPs is regulated by several mechanisms like gene conversion, gene duplication, phase variation and allelic variation <sup>100</sup>.

The blood group antigen binding protein A (BabA, HopS) is the best characterized *H. pylori* adhesin and binds to fucosylated blood group antigen Lewis b <sup>118</sup> present in the normal gastric mucosa. Additionally, BabA adheres to the salivary mucin MUC5B, but also recognizes related terminal fucose residues found on blood group antigens on the surface of mucins like MUC5B <sup>119</sup>. BabA-mediated adherence on the epithelial surface is a potentiator of the *H. pylori* T4SS secretion system activity, thus triggering the transcription of genes enhancing inflammation, development of intestinal metaplasia and associated precancerous transformations <sup>120</sup>. BabA binding is also important for the induction of DNA double-strand breaks and the DNA damage response in host cells <sup>121</sup>. The expression of the *babA* gene is strictly regulated. By RecA-dependent intragenomic recombination, gene conversion between the *babA* and the paralogous *babB* gene occurs and leads to the generation of chimeric genes (*babB/A* and *babA/B*) that may have lost or gained Lewis b binding ability. Furthermore, phase variation may generate mixed genotypes of BabA positive and negative bacteria through varying numbers of cytidine-thymidine (CT) repeats in the 5' region of the gene and slipped strand mispairing (SSM) <sup>122</sup>.

Upon colonization, persistent infection induces an inflammatory response with concomitant expression of sialylated antigens which are recognized by the sialic acid binding adhesin A (SabA, HopP). In persistent infection and chronically inflamed tissue, SabA has been shown to adhere to sialyl-Lewis x and a antigens <sup>123</sup>, the salivary mucins MUC7 and MUC5B and sialylated structures found on the extracellular matrix protein laminin as well as on erythrocytes and neutrophils <sup>119 124 125</sup>. Like BabA, SabA is strongly regulated by gene conversion between *sabA* and the paralogous genes *sabB* and *hopQ* <sup>126</sup> and phase variation by SSM both in the 5' region of the gene and in the *sabA* promoter <sup>127</sup>.

Since not all *H. pylori* strains produce BabA and SabA, further bacterial proteins must be involved in adhesion of bacteria to gastric mucins and cells. HopD (N,N'-diacetyllactosediamine (lacdiNAc)-binding adhesin (LabA)), a further member of the putative adhesin branch of the Hop family, was recently found to recognize lacdiNAc carried by gastric mucins <sup>128</sup>. The adherence associated lipoproteins AlpA (HopC) and AlpB (HopB) have been discovered to bind to laminin and to be involved in gastric colonization <sup>129</sup>. The receptors for the adhesins HopZ, OipA (outer inflammatory protein A) and HopQ, have not been discovered yet. Besides these OMPs already characterized as adhesins, further proteins may be involved in the attachment to host cells suggesting multiple or variable modes of host-pathogen interactions adapted to the unique gastric environment. Most of the bacteria are motile at the surface of the mucus layer and only a

small percentage is actually adhering to the gastric epithelium. The adherent state is favored due to the constant release of nutrients of the *H. pylori* infected and damaged host cells. The motile state reduces the risk of elimination by the host's immune responses <sup>100</sup>.

## 4 Aims of this thesis

It has already been reported for a set of pathogens like *N. gonorrhoeae* and *N. meningitidis*, *M. catarrhalis* and *E. coli* that they exploit members of the CEACAM family as entry site in their host. CEACAMs are involved in main cellular functions like cell-cell adhesion, homeostasis and regulation of the immune system which makes them a favorable target regarding their central role in cells. Especially CEACAM1 is reported to modulate both, the innate and the adaptive immunity. In addition to that, CEACAM1, CEACAM5 and CEACAM6 are involved in tumorigenesis.

The epithelial members CEACAM1, 5 and 6 together with the granulocyte receptor CEACAM3 are exploited by Gram-negative and human restricted pathogens. The basic aim of this thesis was to find out, whether *H. pylori* interferes with members of the CEACAM family. A bacterial pull-down assay already used for *M. catarrhalis*, *N. gonorrhoeae* and *N. meningitidis* should be established to analyze the capability of *H. pylori* to bind to the N-terminal domain of CEACAMs.

In the attempt to identify the bacterial adhesin interacting with CEACAMs, several proteins of the Hop family should be targeted, since members of this family have already been identified before as adhesins, but the corresponding ligands were not discovered yet. Deletion mutants of BabA, BabB, SabB, HopZ and HopQ were generated and analyzed for a loss or impairment in CEACAM-binding.

Since one of the main reasons for adherence of *H. pylori* to the gastric epithelium is the translocation of the effector protein CagA into host cells to modulate cellular processes, the bacterial binding to CEACAMs was also analyzed for its impact on CagA translocation. It was also planned to evaluate the effect of the generated OMP mutants on the CagA translocation capability. Therefore, CagA translocation was studied by *in vitro* infections of gastric cancer cells with wildtype and adhesion-defective mutant *H. pylori* strains.

# Results

---

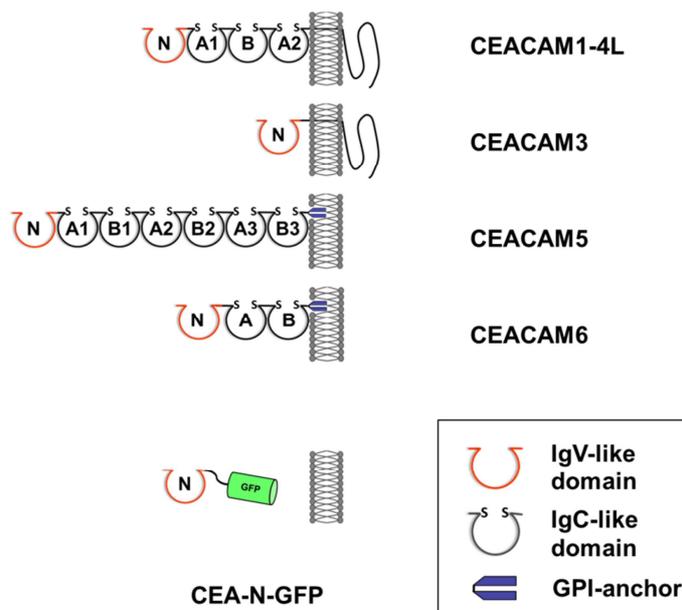
## 1 *H. pylori* binds to CEACAM-N-GFP constructs

### 1.1 ESTABLISHMENT OF A BACTERIAL PULL-DOWN ASSAY

*M. catarrhalis*, *N. meningitidis*, *N. gonorrhoeae* and other Gram-negative bacteria are known to bind to members of the CEACAM family. It could be demonstrated for CEACAM1, 5 and 6 that pathogenic bacteria use these epithelial CEACAMs as entry receptors for their host cells <sup>130</sup>. Additionally, the granulocyte receptor CEACAM3 was discovered as pathogen receptor. Human CEACAM3 is part of the host defense mechanism of the innate immune system against pathogens. Receptor-bound bacteria are internalized and phagocytosed <sup>44</sup>. The four pathogen interacting receptors CEACAM1, 3, 5 and 6 represent therefore the most probable CEACAM candidates to examine for a possible interaction with *H. pylori* (Figure 9).

To test whether *H. pylori* can bind members of the CEACAM family, a bacterial pull-down assay was established that has already been used for *M. catarrhalis* and *Neisseria* species. Using this bacterial pull-down assay, it was clearly shown that the N-terminal, IgV-like domain of CEACAMs was sufficient for binding <sup>131</sup>.

Constructs for CEACAM1, 3, 4, 5, 6, 7 and 8 for testing the interaction with *H. pylori* strains were obtained from Christof Hauck (Konstanz). The N-terminal domain of the distinct CEACAMs was fused to GFP (further described as CEA-N-GFP constructs) and transfected into HEK293 cells. After 48 h, the supernatant, that contained the soluble CEA-N-GFP constructs, was harvested and used in the pull-down assay. A sample of bacteria alone and supernatant of HEK293 cells transfected with GFP alone served as negative controls.



**Figure 9: Human CEACAMs 1, 3, 5 and 6.** Four members of the CEACAM family are depicted. Each member has a distinct variable IgV-like domain (red) and different numbers of IgC-like domains that can be distinguished in A and B subsets (black).

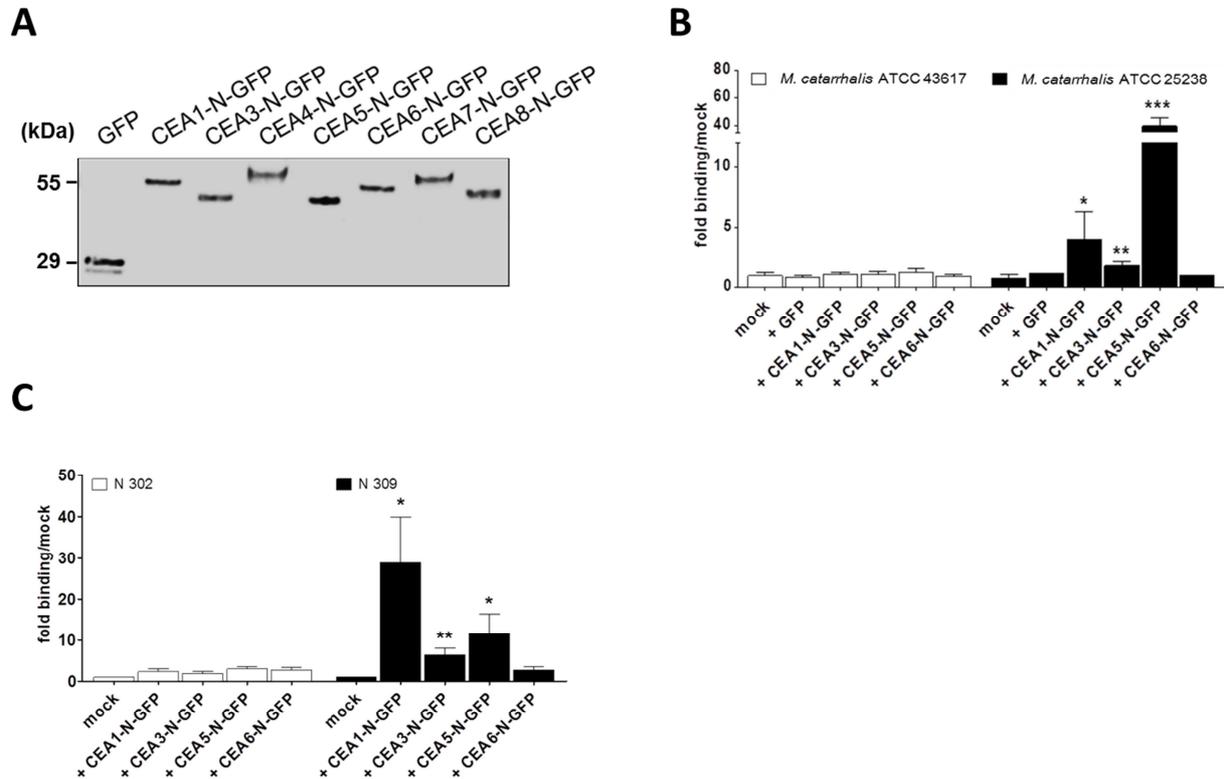
The cytoplasmic domain of the CEACAM1-4L-isoform has two ITIM motifs, whereas the cytoplasmic domain of CEACAM3 contains an ITAM motif.

CEACAM5 and 6 are linked to the cytoplasmic membrane by a GPI-anchor. For the soluble CEA-N-GFP fusion proteins the N-terminal IgV-like binding domain was fused to GFP. The constructs were transfected into HEK293 cells and after 48h the supernatant was harvested and used in the bacterial pull-down assay.

Soluble CEA1-N-GFP – CEA8-N-GFP and the negative control GFP were analyzed on a SDS gel to check the constructs. The molecular weight of the fusion proteins was around 55 kDa, except for GFP which had a molecular weight of 29 kDa. The actual size of the constructs correlated with the expected size (Figure 10 A) <sup>131</sup>. In each pull-down assay, equal amounts of fusion proteins were used to make sure that the binding capability of bacteria to each construct was comparable.

The GFP fluorescence signal of the bacteria was detected by flow cytometry as mean fluorescence intensity (MFI) value. If the bacterium had bound the construct, the flow cytometer detected in the FITC channel a fluorescence signal shift to the right, which correlated with higher MFI values compared to bacterial controls. No shift would be observed if the bacterium has not been binding to the construct. The MFI values of the constructs were normalized to the mock control, which consisted of a sample with only bacteria and was set to the value of 1. The graphs show the fold binding of bacteria to CEA-N-GFP constructs compared to the negative control. *M. catarrhalis* and *N. gonorrhoeae* strains were used as controls for bacterial binding <sup>43</sup>. *M. catarrhalis* strain ATCC 43617 was described to be negative for binding to CEACAM members. *M. catarrhalis* strain ATCC 25238 was used in the experiments as positive control, which showed significant binding for CEA1-N-GFP (3.9 fold), CEA3-N-GFP (1.7 fold) and CEA5-N-GFP (38.8 fold) compared to the mock control, but not for other constructs (Figure 10 B). *N. gonorrhoeae* strain N302 was negative for CEACAM-binding and *N. gonorrhoeae* strain N309 showed

significant binding for CEA1-N-GFP (28.9 fold), CEA3-N-GFP (6.3 fold) and CEA5-N-GFP (11.7 fold), respectively (Figure 10 C).



**Figure 10: Controls for the bacterial pull-down assay.**

(A) Equal amounts of CEA-N-GFP constructs were used in the bacterial pull-down assay. All constructs were put on a SDS gel and checked for the specified size. (B) *M. catarrhalis* strains ATCC 43617 and ATCC 25238 as well as (C) *N. gonorrhoeae* strains N302 and N309 were tested for the interaction with CEA-N-GFP constructs to establish the bacterial pull-down assay. The GFP fluorescence signal of the bacteria was detected by flow cytometry as MFI value. If the bacterium had bound the construct, the flow cytometer detected a fluorescence signal shift to the right, which correlated with higher MFI values compared to the negative control. The fold binding was determined compared to the sample containing only bacteria which was set to the value of 1. The strains ATCC 43617 and N302 served as negative controls for CEACAM-binding and the strains ATCC 25238 and N309 as positive controls.

Student's T-test; \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ . Values are means +/- SD, n = 3

## 1.2 *H. PYLORI* STRAIN P12 BINDS CEA1-N-GFP AND CEA5-N-GFP

To address the question whether *H. pylori* can attach to CEACAMs, the *H. pylori* strain P12 was tested in the pull-down assay. P12 showed significant binding to CEA1-N-GFP and CEA5-N-GFP, but to none of the other tested constructs. The binding to CEA1-N-GFP was 24.5 fold higher than the binding to the mock control. The interaction with CEA5-N-GFP was even stronger with a mean value of 50.8 fold binding compared to the mock control (Figure 11 A).

Interestingly, the complete bacterial population was able to interact with CEA1-N-GFP (91% of P12 population) and CEA5-N-GFP (95% of P12 population). The constructs CEA3-N-GFP, CEA4-N-GFP, CEA6-N-GFP, CEA7-N-GFP and CEA8-N-GFP showed identical MFI values as the negative controls, mock and GFP, thus no significant interaction could be found between P12 and these constructs (Figure 11 B).

The analysis of pull-down samples by western blotting using an  $\alpha$ -GFP antibody is an even more sensitive method than analysis by flow cytometry, since weak interactions between bacteria and constructs can be traced <sup>131</sup>. Therefore, all samples have been analyzed again using the western blot method to test the negative constructs.

The western blot results demonstrated that P12 did not bind to CEA3-N-GFP, CEA4-N-GFP, CEA6-N-GFP, CEA7-N-GFP and CEA8-N-GFP in the bacterial pull-down assay (Figure 11 C), thus confirming the results obtained by flow cytometry.

P12 was found in the assays to bind CEA1-N-GFP considerably weaker than CEA5-N-GFP. By western blot analysis, the exact ratio was determined. The detected band intensities of CEA1-N-GFP and CEA5-N-GFP, respectively, were normalized against the detected band intensities for the bacterial protein RecA, to compare the affinity of P12 for the two constructs. The densitometric evaluation of three independent pull-down assay experiments in the western blot showed for P12 and CEA5-N-GFP a 4.4 times higher affinity in binding as compared to P12 and CEA1-N-GFP that was set to the value of 1 (Figure 11 D).

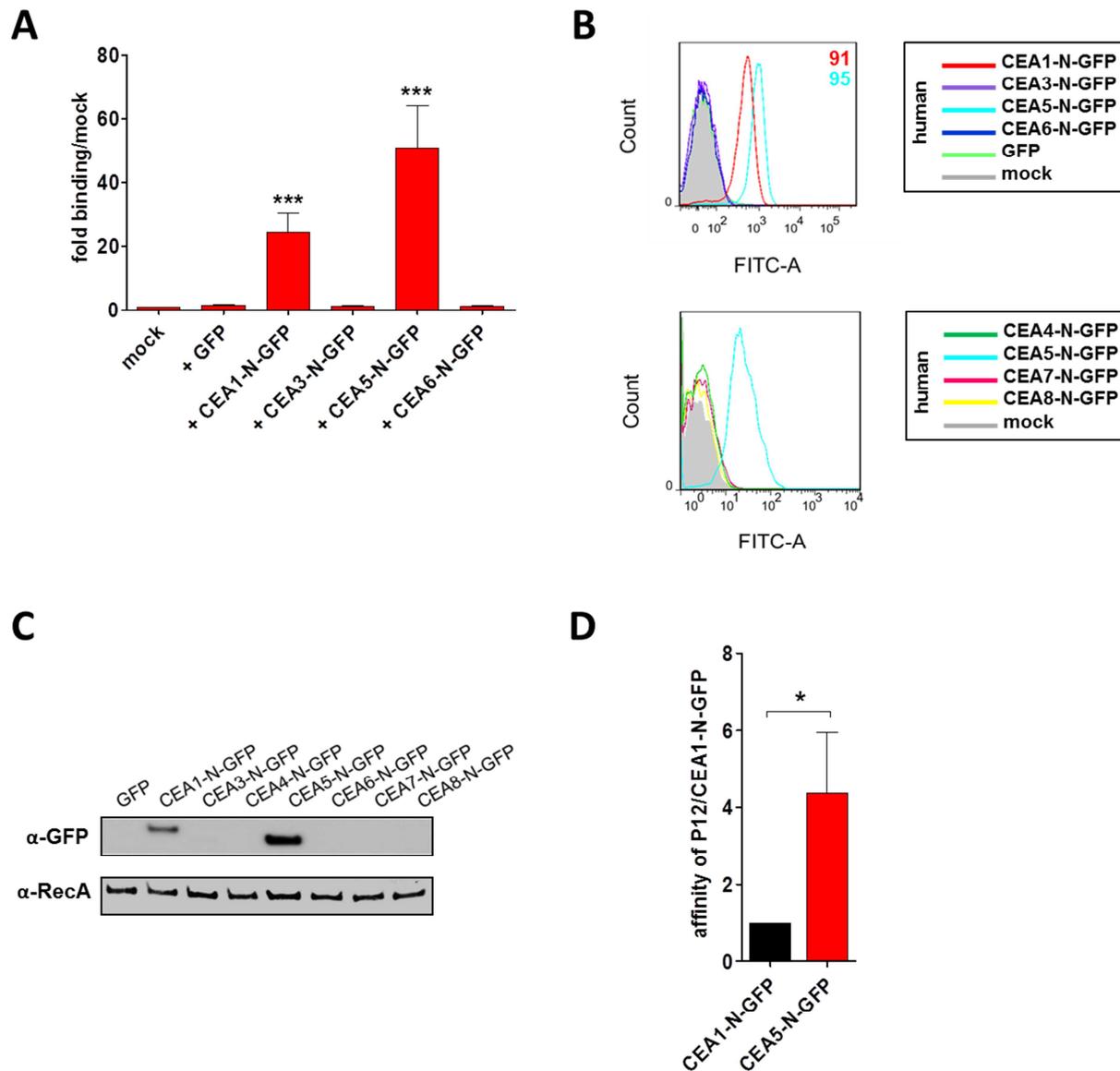


Figure 11: *H. pylori* strain P12 is binding to CEA1-N-GFP and CEA5-N-GFP.

(A) The flow cytometer analysis of P12 binding to CEA1-N-GFP, CEA3-N-GFP, CEA5-N-GFP and CEA6-N-GFP constructs is depicted. The MFI values were determined and compared to the mock control. (B) The flow cytometer readout of the P12 pull-down assay with CEA1-N-GFP, CEA3-N-GFP, CEA5-N-GFP and CEA6-N-GFP (top) as well as CEA4-N-GFP, CEA5-N-GFP, CEA7-N-GFP and CEA8-N-GFP (bottom) is shown. The P12 population that was negative for the interaction with constructs did not shift to the right (grey area). Bacteria that bound to the constructs shifted to the right. (B, top) The numbers represents the percentage of the P12 population that attached to the CEA-N-GFP constructs. (C) The pull-down assay samples were evaluated by western blot analysis. The samples were loaded on a gel and blotted against an  $\alpha$ -GFP antibody. The detected band intensities were compared to the bacterial protein RecA. (D) The detected band intensities of CEA1-N-GFP and CEA5-N-GFP, respectively, were normalized against the band intensities for the bacterial protein RecA to compare the affinity of P12 for the two constructs (three independent pull-down assay experiments).

Student's T-test; \* $p \leq 0.05$ , \*\*\* $p \leq 0.001$ . Values are means +/- SD,  $n = 3$

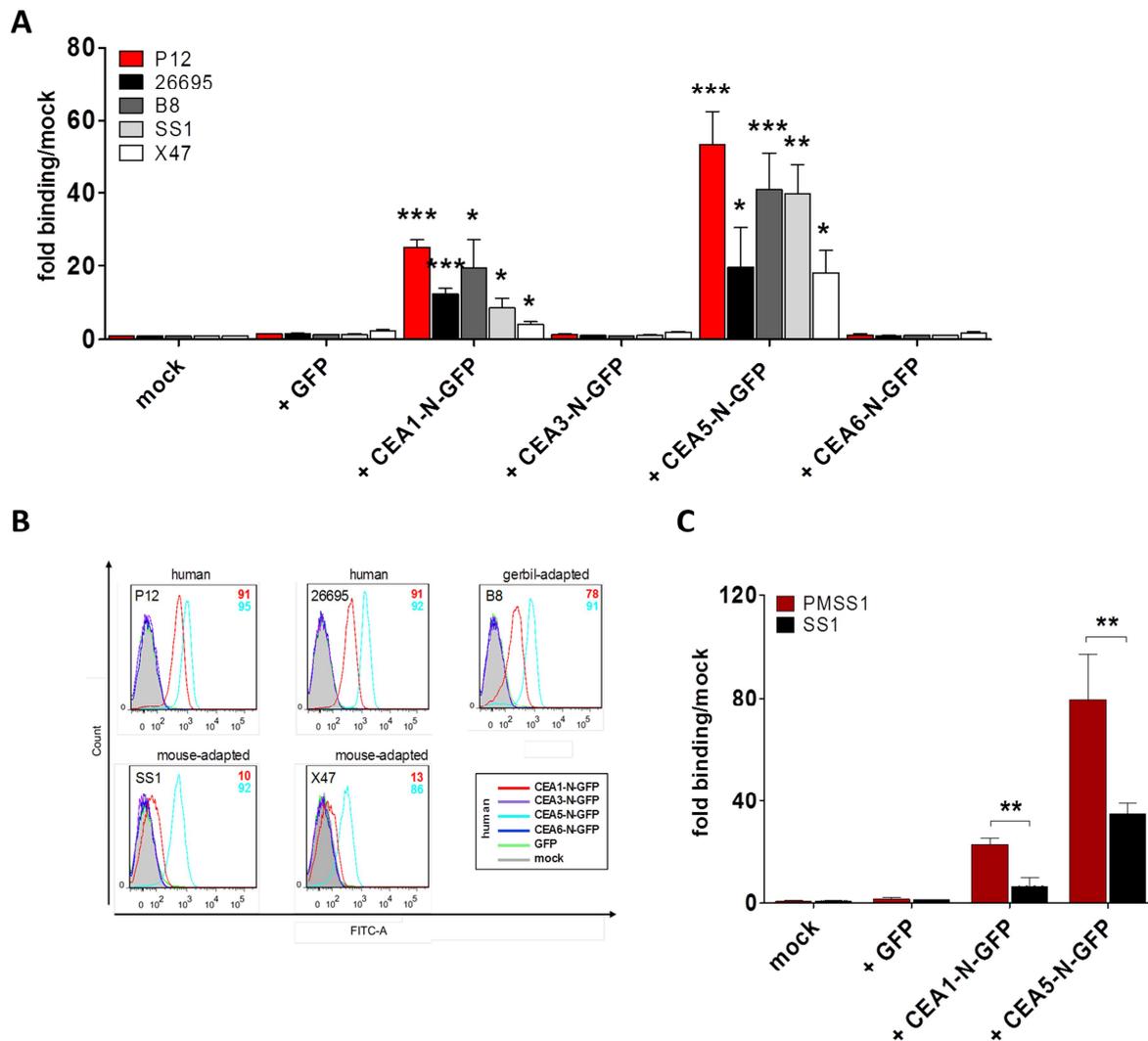
### 1.3 CEACAM-BINDING OF *H. PYLORI* STRAINS

Since bacterial pull-down assay experiments revealed a clear binding capability of *H. pylori* strain P12 for CEA-N-GFP constructs, additional strains were evaluated to determine the extent of the binding in the highly diverse *H. pylori* group. Besides P12, the strains 26695, B8, X47 and SS1 were analyzed for CEACAM-binding. Like P12, strain 26695 also belongs to the *hpEurope* group and its genome is fully sequenced and published. The *hpEurope* strain B8 was analyzed, since it was interesting to characterize a *H. pylori* strain adapted to the animal model of Mongolian gerbils (*Meriones unguiculatus*). Although the genome of SS1 (Sydney strain 1) and X47 is only known partially, the strains were chosen, due to the fact that they are adapted to the mouse model. A further characteristic is that they are negative for cagPAI in contrast to P12, 26695 and B8.

The experimental setting of P12 was used in the pull-down assays and the samples were analyzed by flow cytometry. The results of the four strains showed a binding capability to CEACAM1 and CEACAM5, which nicely confirmed the results obtained for P12. Furthermore, the strains were negative for binding to the constructs CEA3-N-GFP, CEA4-N-GFP, CEA6-N-GFP, CEA7-N-GFP and CEA8-N-GFP, as seen for P12.

A closer look at the results revealed that the fold binding for the single strains varied in its degree. The strain 26695 was binding to CEACAM1 and CEACAM5 with 12.2 and 19.5 fold the mock control, respectively, which was about half the binding capability of P12. The gerbil adapted strain B8 had mean values of 19.3 fold binding to CEACAM1 and 40.9 fold binding to CEACAM5 similar to P12. The mouse adapted strain SS1 showed similar binding to CEACAM5 (39.8 fold) as P12 and B8, but the CEACAM1 interaction was considerably lowered to 8.5 fold binding of the negative control. The mouse strain X47 was the strain that bound both constructs the least. The interaction with CEACAM1 was at 4.0 fold binding of the mock control, CEACAM5 matched strain 26695 with a mean of 18.0 fold binding (Figure 12 A).

All strains were able to generally bind to CEACAM1 and 5 in the pull-down assay. Nearly the whole population of strains P12, 26695, B8 and SS1 bound to CEA5-N-GFP (91%– 95% of the respective population) with similar binding intensity. For the strain X47, the binding intensity was slightly reduced, but represented still the main part of the bacterial population (86%). The results for CEA1-N-GFP were completely different to the obtained results for CEA5-N-GFP. Although P12 and 26695 showed still binding of 91% of the respective population, the interaction with the CEA1-N-GFP construct was significantly reduced in the binding intensity for both the gerbil and mouse adapted strains. Especially the mouse adapted strains had the weakest binding intensity to CEACAM1. A shift of only 10% and 13% of the examined populations was registered for CEA1-N-GFP (Figure 12 B).



**Figure 12: Test of further *H. pylori* strains for their capability to bind CEACAM1 and CEACAM5.**

(A) The strains 26695, B8, SS1 and X47 were analyzed for their binding to CEACAMs in the bacterial pull-down assay and compared to P12. (B) The flow cytometer readout of the single strains is depicted. Bacteria that bound to the constructs shifted to the right. The percentage of the single strains' population that shifted to the right is revealed by the numbers on the top right in each flow cytometer readout. (C) PMSS1 is the original clinical isolate that was passaged through the mouse which resulted in the reisolated mouse strain SS1. The two strains that differed by animal passage of SS1 were compared in the bacterial pull-down assay and analyzed for a change in the CEACAM receptor interaction.

Student's T-test; \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ . Values are means +/- SD,  $n = 3$

Additionally to the five strains that were analyzed so far for CEACAM-binding, a further strain was tested in the pull-down assay. *H. pylori* strain PMSS1 (pre-mouse Sydney strain 1, also called 10700) was analyzed to compare the binding to CEACAMs directly to the closely related strain SS1. PMSS1 is the original clinical isolate that was passaged through the mouse which resulted in the reisolated mouse strain SS1<sup>132</sup>. The two strains that differed only by animal passages of SS1 were compared in the bacterial pull-down assay and analyzed for any change in CEACAM receptor interaction. Interestingly, the binding to CEACAMs differed significantly in the two strains. PMSS1 bound both constructs about three times better than SS1. Compared to the mock control, PMSS1 showed 22.9 fold binding for CEA1-N-GFP, whereas the fold binding of SS1 was reduced to the mean value of 6.4. For the CEA5-N-GFP construct, PMSS1 was interacting with 79.9 fold binding even stronger than the P12 strain that showed so far the highest binding to CEA5-N-GFP. In contrast to PMSS1, SS1 strain was significantly impaired at a level of 34.7 fold binding to CEA5-N-GFP (Figure 12 C).

## 1.4 SPECIFICITY OF CEACAM-BINDING

All tested *H. pylori* strains were able to bind to CEA1-N-GFP and CEA5-N-GFP in the bacterial pull-down assay. Next, it was examined, if this binding ability to CEACAM receptors was restricted to human-specific *Helicobacter* only. Since there are orthologous CEACAMs found in various animal species, it was promising to test binding of *H. pylori* to these as well.

So far, only *Helicobacter* of the species *pylori* were tested. Therefore, non-*pylori* *Helicobacter* species, *Helicobacter nemestrinae* and *Helicobacter mustelae*, were examined for their ability to bind to CEACAMs. *H. nemestrinae* was isolated from the gastric mucosa of pigtailed macaque and is the closest related species to *H. pylori*<sup>133</sup>. *H. mustelae* colonizes ferrets and, besides *H. pylori*, is the only known *Helicobacter* species that can cause gastric ulceration and cancer in its host<sup>134</sup>. The non-*pylori* *Helicobacter* were tested in the standardized pull-down assay and the evaluation of the CEACAM-binding ability was done by flow cytometry. Both *H. nemestrinae* and *H. mustelae* could not interact with CEA1-N-GFP and CEA5-N-GFP in the bacterial pull-down compared to P12 (Figure 13 A).

Orthologues of CEACAM1, the primordial member of the CEACAM family, are found in numerous species such as mouse, rat or cattle<sup>80 75</sup>. In contrast to that, CEACAM5 is rarely described in other species than humans. The only orthologue has been discovered in the rhesus macaque<sup>52</sup>. Since there are also strains of *H. pylori* that were isolated from mice or rhesus macaques, the question was whether *H. pylori* can bind to the orthologous CEACAMs found in these species. To address this, CEA1-N-GFP constructs of mouse (CEACAM1<sup>a</sup> isoform), cattle and dog and the CEA5-N-GFP construct of the rhesus macaque along with human constructs of CEACAM1 and CEACAM5 were tested for binding with the *H. pylori* strains P12, 26695, B8, SS1 and X47 in the standardized pull-down assay. Compared to the human CEA1-N-GFP and CEA5-N-GFP constructs, the tested strains showed neither binding to the orthologous CEACAM1 variants nor the macaque construct of CEACAM5 in the pull-down assay. Furthermore, the hypothesis that the mouse adapted strains SS1 or X47 might be able to bind to the murine CEA1-N-GFP construct could not be confirmed (Figure 13 B).

Taken together, the results suggest that only the genus *H. pylori* binds CEACAMs, but not the closely related other genera *H. nemestrinae* and *H. mustelae*. Furthermore, *H. pylori* interaction with CEACAMs was restricted to the human receptor constructs in the pull-down assay.

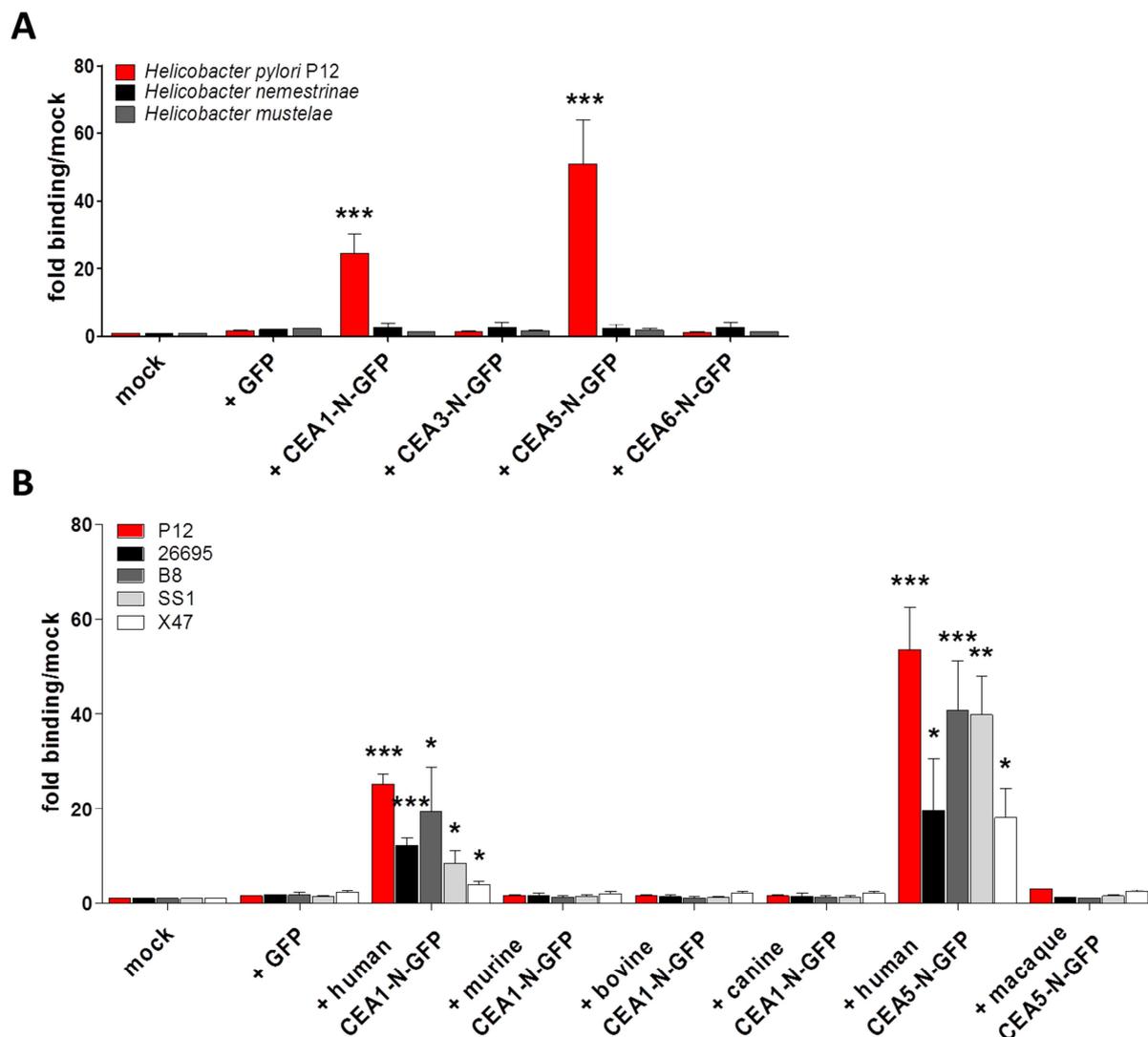


Figure 13: Exclusive and specific binding of the genus *H. pylori* to human CEACAMs.

(A) The non-*pylori* *Helicobacter* species, *H. nemestrinae* and *H. mustelae*, were tested in the pull-down assay for binding to human CEA-N-GFP constructs 1, 3, 5 and 6.

(B) *H. pylori* strains P12, 26695, B8, SS1 and X47 were tested in the pull-down assay for interaction with orthologous constructs of CEA1-N-GFP from mouse, cattle and dog or CEA5-N-GFP from rhesus macaque.

Student's T-test; \* $p \leq 0.05$ , \*\* $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ . Values are means +/- SD, n = 3

## 2 HopQ is the adhesin binding to CEACAM1 and CEACAM5

### 2.1 GENERATION OF MARKERFREE *H. PYLORI* OMP MUTANTS

The bacterial pull-down assay revealed that *H. pylori* binds significantly to the human receptors CEACAM1 and CEACAM5. To find out which adhesin is interacting with the receptors, several promising *H. pylori* outer membrane proteins were deleted by the streptomycin contraselection strategy and the resulting mutants were then analyzed for a loss or impairment in CEACAM-binding.

The streptomycin contraselection strategy is based on the fact that the streptomycin-sensitive (*rpsL<sup>S</sup>*) allele is dominant over the resistance-conferring mutant allele (*rpsL<sup>R</sup>*)<sup>135</sup>. In the first step, a plasmid consisting of a two-gene cassette with an *erm<sup>R</sup>* resistance gene, the dominant streptomycin susceptibility gene *rpsL<sup>S</sup>* of *Campylobacter jejuni* and the flanking up- and downstream regions of the gene of interest was transformed into a streptomycin-resistant *H. pylori* strain. By homologous recombination and selection for erythromycin, the gene to be deleted was replaced by the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette. In a second step, a plasmid containing only the flanking up- and downstream regions of the gene of interest was introduced into the *H. pylori* transformants with the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette. The *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette is replaced by homologous recombination and selection for streptomycin. The strain had now a deleted gene of interest without a resistance marker in the locus of the gene of interest. Additionally, the strain is again streptomycin-resistant, which allows the repeated use of this method for the deletion of further genes (Figure 14 A).

A variation of this method represents the “Xer-cise” method, a one-step transformation technique<sup>136</sup> that makes use of the *H. pylori* Xer recombinase which cuts DNA at *difH* sites in the genome. Instead of the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette cloned between the up- and downstream regions of the gene of interest, a plasmid containing a *difH - rpsL<sup>S</sup>-cat<sup>R</sup> - difH* -cassette was transformed into a streptomycin-resistant wildtype strain. By homologous recombination and selection for chloramphenicol, transformants had the gene of interest replaced by the *difH - rpsL<sup>S</sup>-cat<sup>R</sup> - difH* -cassette. Immediately after appearance of the first colonies, the positive transformants were passaged onto a streptomycin-resistant plate. The clones growing on the plate were transformants that had cut out the resistance cassette by the Xer-recombinase. This method generates a markerfree deletion mutant with one transformation.

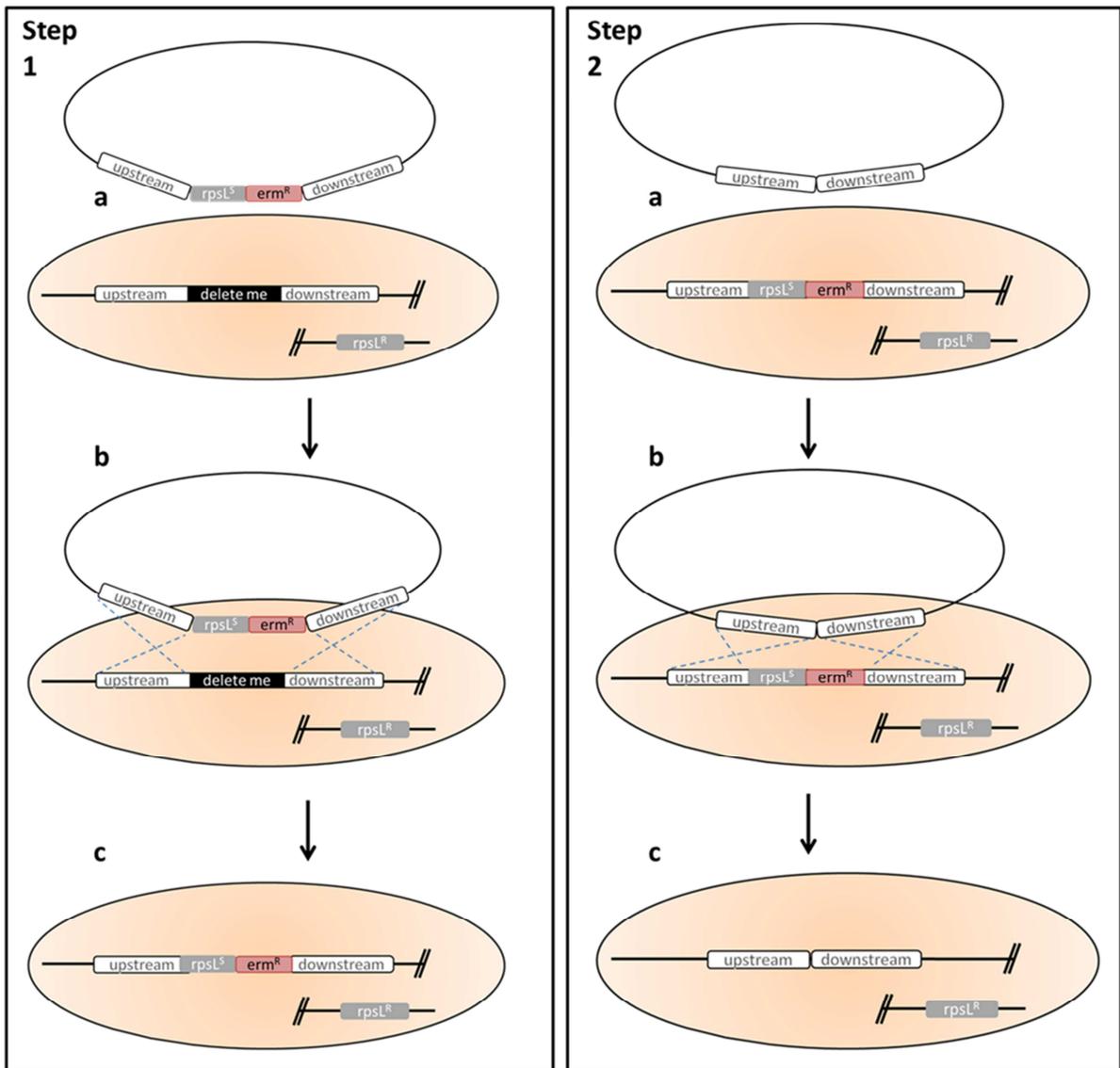
Mutants were verified for the successful deletion of the gene of interest by polymerase chain reaction (PCR): primers were designed that bound in the up- and downstream region close to the gene of interest. The genomic DNA was amplified with the primers and the product was correlated in size to three different stages. Figure 14 B shows exemplarily the results for the

deletion of the gene *hopZ* with the obtained three different fragments (1, 2, 3) corresponding to three stages: 1. The wildtype strain: the gene of interest is still in the original locus which gives the largest fragment; 2. The transformants with the resistance cassette: the gene of interest is replaced by the *rpsLS-erm<sup>R</sup>*-cassette; 3. The markerfree deletion mutant: the *rpsLS-erm<sup>R</sup>*-cassette has been cut out which leads to a markerfree mutant.

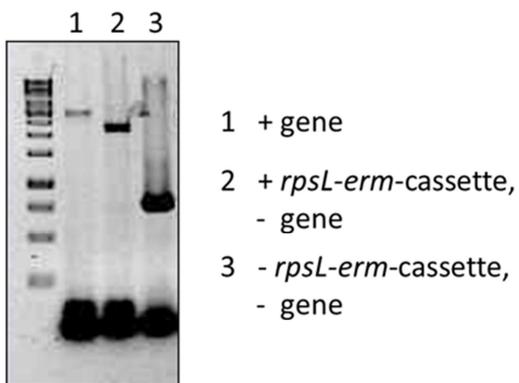
The OMPs to be deleted belonged to the putative adhesin branch of the Hop family <sup>137</sup>. The branch consists of 11 proteins that share a strong sequence homology with each other and include already known adhesins like BabA <sup>118</sup>, SabA <sup>123</sup> and HopZ <sup>138</sup> (Figure 14 C). The OMPs were examined for interaction with CEACAM1 and/or CEACAM5.

For the generation of OMP mutants, the strain P12 was chosen, since the complete genome sequence is published <sup>139</sup> and the strain showed strong binding to CEACAMs. In the special case of P12, instead of one *sabA* and one *sabB* gene, two identical copies of *sabB* are found in the *sabA* and *sabB* gene locus. Furthermore, the strain has no *babC*, but two identical copies of *babB*. In order to create a SabB and a BabB deletion mutant, respectively, two genes had to be deleted instead of one. The OMP deletion mutants generated for P12 were P12 $\Delta$ *babA*, P12 $\Delta$ *babB*, P12 $\Delta$ *sabB*, P12 $\Delta$ *hopZ* and P12 $\Delta$ *hopQ*.

**A**



**B**



**C**

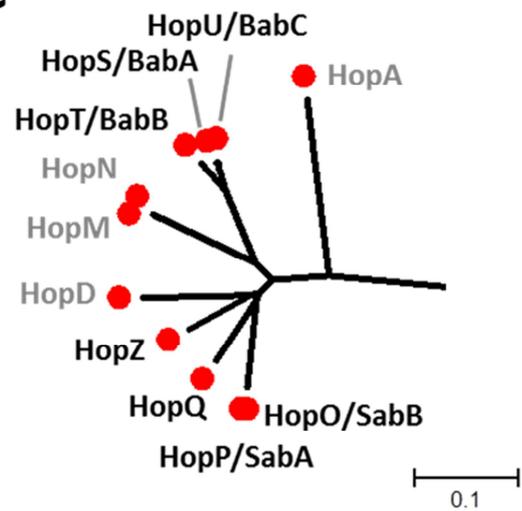


Figure 14: Generation of markerfree outer membrane protein mutants of *H. pylori*.

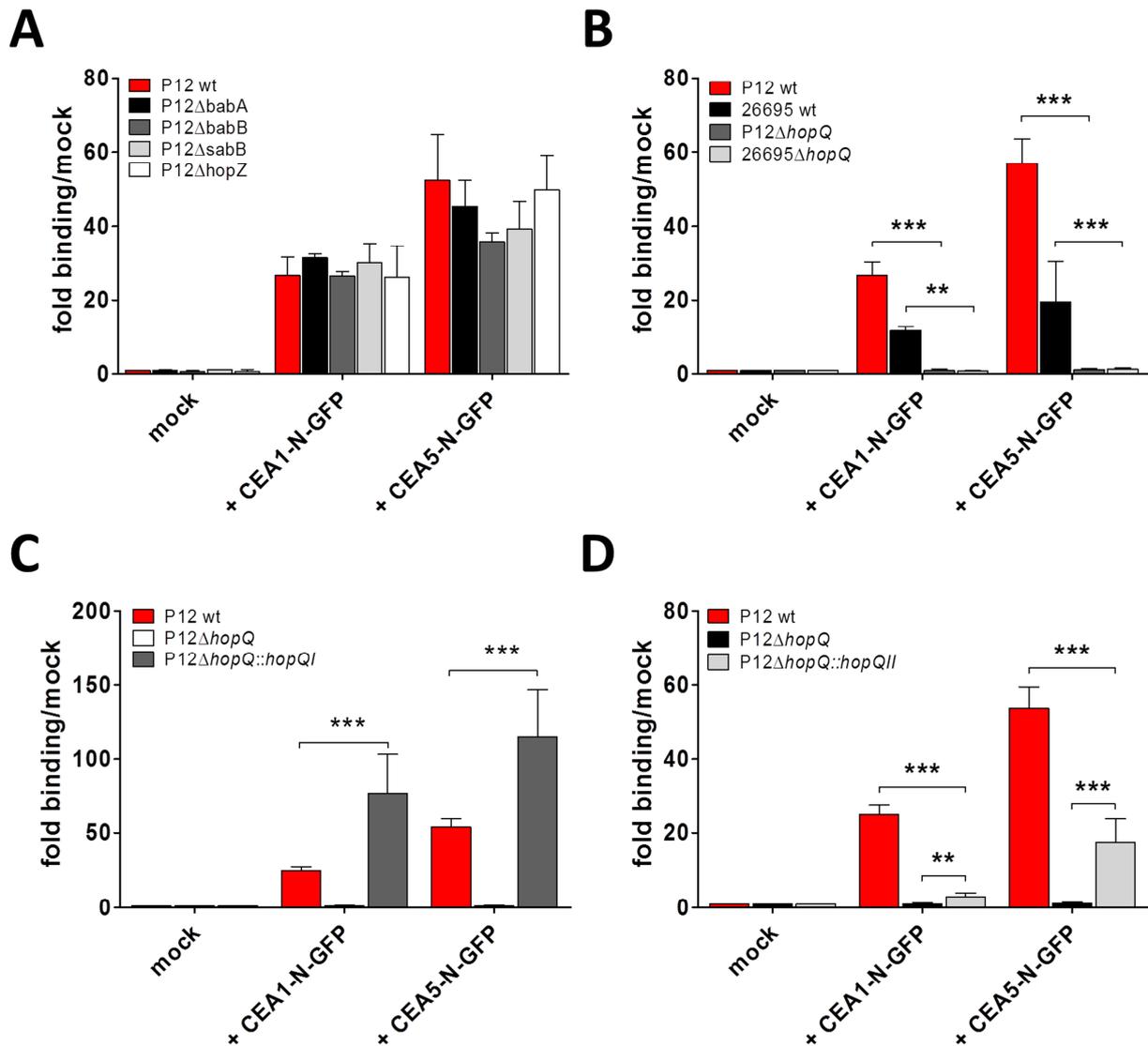
(A) Step1a: The markerfree gene deletion mutants are generated by streptomycin contraselection. A streptomycin-resistant wildtype strain was transformed with a plasmid consisting of a two-gene cassette with an *erm<sup>R</sup>* resistance gene, the dominant streptomycin susceptibility gene *rpsL<sup>S</sup>* of *C. jejuni* and the flanking up- and downstream regions of the gene of interest was transformed into a streptomycin-resistant *H. pylori* strain. Step1b: By homologous recombination, the gene of interest was replaced by the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette. Step1c: A screen for positive transformants was performed by selection for erythromycin. Step2a: Erythromycin resistant and streptomycin sensitive clones were transformed with a plasmid containing only the flanking up- and downstream regions of the gene of interest. Step2b: By homologous recombination, the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette was cut out. Step2c: By screening for streptomycin-resistant transformants, clones were selected with a deletion for the gene of interest and a removed selection marker. (B) The markerfree deletion of a gene of interest was verified by PCR reaction and the amplified samples loaded onto an agarose gel: primers binding in the up-and downstream region close to the gene of interest were amplifying distinct fragments in size. By the size of the amplified fragment, it was determined, if the gene was still in the genome (1), if the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette had replaced the gene of interest (2) or if the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette was cut out and the gene deleted without remaining markers (3). (C) The putative adhesin branch of the Hop family of proteins. The OMP deletion mutants generated in this work for P12 were P12 $\Delta$ *babA*, P12 $\Delta$ *babB*, P12 $\Delta$ *sabB*, P12 $\Delta$ *hopZ* and P12 $\Delta$ *hopQ* and are depicted in black.

## 2.2 *H. PYLORI* BINDING IS HOPQ DEPENDENT

The generated OMP mutants P12 $\Delta$ *babA*, P12 $\Delta$ *babB*, P12 $\Delta$ *sabB*, P12 $\Delta$ *hopZ* and P12 $\Delta$ *hopQ* were examined in the bacterial pull-down assay for a loss or a reduced interaction with CEACAM1 and CEACAM5, respectively. Although it has been established that BabA binds Lewis b antigens<sup>118</sup> and SabA is interacting with sialyl-Lewis x and a antigens<sup>124</sup>, they might have a potential influence on CEACAM-binding. So far, no receptor is described for the three generated deletion mutants SabB, HopZ and HopQ. HopQ was recently identified as novel T4SS virulence factor<sup>140</sup>.

The deletion mutants were compared to the binding of P12 wildtype to CEACAM1 and CEACAM5. The deletion mutants P12 $\Delta$ *babA*, P12 $\Delta$ *babB*, P12 $\Delta$ *sabB* and P12 $\Delta$ *hopZ* showed no significant difference in binding to CEA1-N-GFP and CEA5-N-GFP compared to the P12 wildtype (Figure 15 A). In strong contrast to P12 $\Delta$ *babA*, P12 $\Delta$ *babB*, P12 $\Delta$ *sabB* and P12 $\Delta$ *hopZ*, the P12 $\Delta$ *hopQ* deletion mutant lost the binding ability to both CEA1-N-GFP and to CEA5-N-GFP completely (Figure 15 B). The level of interaction with CEACAMs was for both constructs reduced to the level of the mock control. To validate the loss of binding to CEACAMs in the P12 $\Delta$ *hopQ* deletion strain, a second  $\Delta$ *hopQ* deletion strain was analyzed for binding ability. 26695 $\Delta$ *hopQ* was already generated and described in Odenbreit *et al.*, 2002<sup>141</sup>. The results of the bacterial pull-down assay showed for the 26695 $\Delta$ *hopQ* deletion mutant also a complete loss of binding to CEA1-N-GFP and to CEA5-N-GFP compared to the 26695 wildtype.

To prove that HopQ is the adhesin interacting with CEACAM1 and CEACAM5, the P12 $\Delta$ *hopQ* deletion mutant was genetically complemented. The *pHel3* shuttle system was chosen as method for complementation, which represents a highly efficient method to transform plasmids into *H. pylori* strains<sup>142</sup>. The *hopQI* gene of P12 and the *hopQII* gene of Tx30a, respectively, were cloned behind an *alpA* promoter and the resulting *pHel3* shuttle plasmid was introduced into the P12 $\Delta$ *hopQ* deletion mutant by electroporation<sup>143</sup>. The P12 wildtype, the P12 $\Delta$ *hopQ* deletion mutant, the *hopQI* complemented P12 $\Delta$ *hopQ* deletion mutant (P12 $\Delta$ *hopQ*::*hopQI*) and the *hopQII* complemented P12 $\Delta$ *hopQ* deletion mutant (P12 $\Delta$ *hopQ*::*hopQII*) were compared in the standardized bacterial pull-down assay. The results showed that the binding was completely reconstituted for both CEACAM1 and CEACAM5 in the genetically complemented strains compared to the P12 $\Delta$ *hopQ* deletion mutant. Moreover, compared to P12 wildtype, a significant increase in binding could be detected for strain P12 $\Delta$ *hopQ*::*hopQI* (Figure 15 C). The strain P12 $\Delta$ *hopQ*::*hopQII* was also capable of interacting with CEACAM1 and CEACAM5, but the fold binding was significantly decreased in comparison to the P12 wildtype (Figure 15 D).



**Figure 15: HopQ is the adhesin interacting with CEACAM1 and CEACAM5.**

(A) The generated OMP mutants P12 $\Delta$ *babA*, P12 $\Delta$ *babB*, P12 $\Delta$ *sabB* and P12 $\Delta$ *hopZ* were examined for binding to CEA1-N-GFP and to CEA5-N-GFP and compared to the P12 wildtype strain.

(B) The  $\Delta$ *hopQ* deletion mutants of strain P12 and 26695 are depicted. The binding to the CEACAM constructs of the deletion mutants is compared to the respective wildtype strains.

The genetically complemented strains (C) P12 $\Delta$ *hopQ*::*hopQI* and (D) P12 $\Delta$ *hopQ*::*hopQII* were compared to the P12 wildtype and the P12 $\Delta$ *hopQ* deletion mutant for the capability to bind to CEA1-N-GFP and to CEA5-N-GFP.

Student's T-test; \*\* $p \leq 0.01$ , \*\*\* $p \leq 0.001$ . Values are means +/- SD, n = 3

## 2.3 HOPQ VARIATIONS IN DIFFERENT *H. PYLORI* STRAINS

### 2.3.1 SEQUENCE ANALYSIS OF HOPQI AND HOPQII

The bacterial pull-down assay established that HopQ is the adhesin binding to human CEACAM1 and CEACAM5. Without HopQ no binding was registered. When the strains P12, 26695, B8, SS1 and X47 were compared to each other, the fold binding effect was different for every single strain, although equal amounts of bacteria and constructs were used in the pull-down assay (Figure 12). HopQ was studied in more detail to find out, if the variations in binding were due to differences in the aa sequence of the adhesin.

HopQ exists in two allelic forms, type I and type II, which are associated with *H. pylori* type I or type II strains, respectively. The more virulent type I class of strains is defined as positive for CagPAI and the s1 type of VacA and produces a HopQ type I (HopQI). The less virulent type II class of *H. pylori* strains is negative for CagPAI and most likely possesses a s2 type of VacA and a HopQ type II (HopQII) <sup>105</sup>. P12, 26695 and B8 are type I strains, whereas the mouse adapted strains belong to the group of type II strains. The HopQ aa sequences of the strains P12, 26695, B8, X47 and SS1 were aligned with the program ClustalW. The aa consent with P12 is shaded in black, the differences are depicted in white (Figure 16 A).

All strains were nearly identical for the N-terminal 50 aa and the 180 aa of the C-terminal part of HopQ. The *hpEurope* strains P12, 26695 and B8 were mostly conserved throughout the HopQ sequence as were the aligned HopQII strains (SS1 and X47), respectively. When the HopQI and HopQII sequences were compared to each other, they varied largely, especially between aa 60 – 100, aa 130 – 250 and aa 378 – 418, where the HopQ sequences were completely different.

The percentage of sequence identity was very high. Between P12 and 26695 the sequence identity was at 96.6%, P12 and B8 had a value of 94.8%. The HopQII containing strains X47 and SS1 were also closely related in sequence with 96.7% identity. Between HopQI and HopQII strains the percentage of sequence identity decreased significantly. Only 72% of sequence identity were found between the HopQI of P12 and the HopQII of X47 and SS1 (Figure 16 B).

Interestingly, B8 as HopQI strain showed between aa 429 – 471 a nearly 100% identical sequence to the mouse adapted HopQII strains X47 and SS1, but differed from the HopQI strains P12 and 26695.

A

	10	20	30	40	50	60	
P12	MKKTKKTI LLSLTLASSLLHAEDNGVFLSVGYQI GEAVQKVKNADKVQKLSDTYEQLSRL						60
26695	MKKTKKTI LLSLTLAASLLHAEDNGVFLSVGYQI GEAVQKVKNADKVQKLSDTYEQLSRL						60
B8	MKKTKKTI LLSLTLAASLLHAEDNGVFLSVGYQI GEAVQKVKNADKVQKLSDAYEQLSRL						60
SS1	MKKTKKTI LLSLTLASSLLHAEDNGVFLSVGYQI GEAVQKVKNADKVQKLSDAYENLNKI						60
X47	MKKTKKTI LLSLTLAASLLRAEDNGVFLSVGYQI GEAVQKVKNADKVQKLSDAYENLNKL						60
	70	80	90	100	110	120	
P12	LTNDNGTN----SKTSAQAI NQAVNNLNERAKTLAGGTTNSPAYQATLLALRSVLGLWNS						116
26695	LTNDNGTN----SKTSAQAI NQAVNNLNERAKTLAGGTTNSPAYQATLLALRSVLGLWNS						116
B8	LTNDNGTN----SKTSAQAI NQAVNNLNERAKTLAGGTTNSPAYQATLLALRSVLGLWNS						116
SS1	LANHDHSNPEAI NTNSATAI NQAI GNLNANTQNLIDKTDNSPAYQATLLALKSTVGLWNS						120
X47	LANHSHSNPEAI NANSATAI NQAI GNLNKNNTQNLIDKTDNSPAYQATLLALKSTVGLWNS						120
	130	140	150	160	170	180	
P12	MGYAVI CGGYTKSPGENNQKNFHYTDENGGTTI NCGGSTNSNGTHSSNGTNTLKADKNV						176
26695	MGYAVI CGGYTKSPGENNQKDFHYTDENGGTTI NCGGSTNSNGTHSSSGTNTLKADKNV						176
B8	MGYAVI CGGYTKSPGENNQKNFHYTDENGGTTI NCGGSTKSDGTHSSNGTNTLKADKNV						176
SS1	IAYAVI CGGYTDKPNHNI TETFYNQPGQ- NSNSITCG-----SNGLGTLPACKNS						169
X47	IAYAVI CGGYTDKPNHNI TETFYNQPGQ- NSNSITCG-----SNGLGTLPACKNS						169
	190	200	210	220	230	240	
P12	SLSI EQYEKI HESYQI LSKALKQAGLAPLNSKGEKLEAHVTTSKYQQ-----DSQTKTTT						231
26695	SLSI EQYEKI HEAYQI LSKALKQAGLAPLNSKGEKLEAHVTTSKPEN-----NSQTKTTT						231
B8	SLSI EQYEKI HESYQI LSKALKQAGLAPLNSKGEKLEAHVTTSKYQQ-----DNQTKTTT						231
SS1	HLSI ECFATLNKAYQI QAALKQG-LPALSDTKKTVETIKTATNAQNI NVNNNNNNAAD						228
X47	HLSI EQFATLNKAYQI QAALKKG-LPALSNNTTTTVEVTIKTATNAQNI NVNNNNNNAAD						228
	250	260	270	280	290	300	
P12	SVI DTTN----DAQNLLTQAQTI VNTLKDYCPMLIAKSSSGSGGAATNTPSWQTAGGGK						287
26695	SVI DTTN----DAQNLLTQAQTI VNTLKDYCPMLIAKSSSESSGAATNTAPSWQTAGGGK						287
B8	SVI DTTN----DAQNLLTQAQTI VNTLKDYCPMLIAKSSSGSSGQATNTPSWQTAGGGK						287
SS1	ATI ETNKNTYI NDAQNLLTQAQTI I NTLQDNCPMLKKGSSS--GTNGANTPSWQTSAN-Q						284
X47	ATI ETNKNTYI NDAQNLLTQAQTI I NTLQDNCPMLKKGSSSN--GGTNGANTPSWQTSAN-Q						286
	310	320	330	340	350	360	
P12	NSCETF GAEFSAASDMI NNAQKI VQETQQL SANQPKNI TQPHNLNLNTPSSLTALAQKML						347
26695	NSCATF GAEFSAASDMI NNAQKI VQETQQL SANQPKNI TQPHNLNLNTPSSLTALAQKML						347
B8	NSCATF GAEFSAASDMI NNAQKI VQETQQL SANQPKNI TQPHNLNLNTPSSLTALAQKML						347
SS1	NSCSVFGTEFSASDMI SNAQNI VQETQQLNTTPLKSI AQPHNFNLNSPNS-VALAQSML						343
X47	NSCSVFGTEFSASDMI SNAQKI VQETQQLNTTPLKSI AQPHNFNLNSPNS-I ALAQSML						345
	370	380	390	400	410	420	
P12	KNAQSQAELKLANQVESDFNKLSSGHLKDYI GKCDMSAI SSTNMTMQSQKNNWNGCAG						407
26695	KNAQSQAELKLANQVESDFNKLSSGHLKDYI GKCDASAI SSANMTMQNQKNNWNGCAG						407
B8	KNAQSQAELKLANQVESDFNKLSSGHLKDYI GKCDASAI SSANMTMQNQKNNWNGCAG						407
SS1	KNAQSQAAVLKLANQVGNDFNRI STGVLNKNI EECNANASSES-----VSNNTWKGKGCAG						398
X47	KNAQSQAQAVLKLANQVGSDFNRI STGVLNKNI EECGANNSSES-----VSNNTWKGKGCAG						400
	430	440	450	460	470	480	
P12	VEETLTSLSKTSAAFDNNQTPQI NQAQNLANTLI QELGNNPFRNMGMI ASSTTNGALNGL						467
26695	VEETLSSLKTSAAFDNNQTPQI NQAQNLANTLI QELGNNPFRNMGMI ASSTTNGALNGL						467
B8	VEETLASLEKSNASFSSQTPQI NQAETLANAI VQELGHNPFKRVGI I SS-QTNGAMNGL						466
SS1	VKQTLTSLSSNASFSSQTPQI NQAETLANAI VQELGHNPFKRVGI I SS-QTNGAMNGL						457
X47	VKQTLTSLSSNASFSSQTPQI NQAQTLANTLI VQELGHNPFKRVGI I SS-QTNGAMNGL						459

	490	500	510	520	530	540	
P12	GVQVGYKQFFGEKKRWGL RYYGFFDYNHAYI KSNFFNSASDVWTYGVGSDLLFNFI NDKN						527
26695	GVQVGYKQFFGEKKRWGL RYYGFFDYNHAYI KSNFFNSASDVWTYGVGSDLLFNFI NDKN						527
B8	GVQVGYKQFFGEKKRWGL RYYGFFDYNHAYI KSNFFNSASDVWTYGVGSDLLFNFI NDKN						526
SS1	GVQVGYKQFFGEKKRWGL RYYGFFDYNHAYI KSNFFNSASDVWTYGVGSDLLFNFI NDKN						517
X47	GVQVGYKQFFGEKKRWGL RYYGFFDYNHAYI KSNFFNSASDVWTYGVGSDLLFNFI NDKN						519

	550	560	570	580	590	600	
P12	TNFLGKNNQI SFGLFGGI ALAGTSWLSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTN						587
26695	TNFLGKNNQI SVGF FGGI ALAGTSWLSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTN						587
B8	TNFLGKNNQI SFGLFGGI ALAGTSWLSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTN						586
SS1	TNFLGKNNQI SFGLFGGI ALAGTSWLSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTN						577
X47	TNFLGKNNQI SFGLFGGI ALAGTSWLSQFVNLKTI SNVYSAKVNTANFQFLFNLGLRTN						579

	610	620	630	640	650	
P12	LARPKKKDSHHAHQHGI ELGVKI PTI NTNYYSF LDTKLEYRRL YSVYLNYYVFAY					641
26695	LARPKKKDSHHAHQHGMELGVKI PTI NTNYYSF LDTKLEYRRL YSVYLNYYVFAY					641
B8	LARPKKKDSHHAHQHGMELGVKI PTI NTNYYSF LDTKLEYRRL YSVYLNYYVFAY					640
SS1	LARPKKKDSHHAHQHGMELGVKI PTI NTNYYSF LDTKLEYRRL YSVYLNYYVFAY					631
X47	LARPKKKDSHHAHQHGMELGVKI PTI NTNYYSF LDTKLEYRRL YSVYLNYYVFAY					633

## B

	P12	26695	B8	SS1	X47	
P12		96.6	94.8	72.5	72.4	P12
26695			94.2	71.8	71.9	26695
B8				75.4	75.2	B8
SS1					96.7	SS1
X47						X47

Figure 16: ClustalW alignment of HopQI and HopQII.

(A) HopQI aa sequences of P12, 26695 and B8 were aligned to the HopQII sequences of SS1 and X47. The consent of every strain to the P12 sequence was depicted in black shades, differences were shown in white. (B) The percentage of sequence identity between the single strains is depicted.

### 2.3.2 HOPQI PRODUCTION IN STRAINS

The more virulent HopQI strains were analyzed in more detail. To study the production level of HopQI in the strains, a peptide antibody was generated (Genosphere Biotechnologies). The specificity of the antibody was designed against HopQI of strain P12, which was the reference strain in the experiments and showed binding to CEACAM1 and CEACAM5.

To generate an antibody specific for HopQI, but not recognizing HopQII, aligned HopQI and HopQII strains were compared. A peptide of 17 aa was finally chosen in P12 that was conserved in sequence for most HopQI strains and distinctly different to and not found in the aligned HopQII sequences (Figure 17 A, red box). The sequence chosen for the antibody was between aa 211 and 227 of the P12 HopQI aa sequence (SKGEKLEAHVTTSKYQQ) (Figure 17 A, grey). The peptide antibody AK298 was made in rabbit and recognized HopQI very efficiently (70 kDa) (Figure 17 B).

Bacterial lysate of the strains P12, 26695, B8, SS1, X47 and PMSS1 was loaded onto a SDS gel and blotted against the antibody AK298. The HopQI signal was compared to the bacterial protein RecA (40 kDa) that was used as control for equal amounts of bacteria loaded onto the SDS gel.

HopQI of P12 and B8 were recognized from the antibody AK298 with an equal signal in strength. In the strain 26695, only the first 14 aa out of 17 were identical to the sequence of P12 and recognized by the antibody. Only a weak signal is produced for HopQI in the blot. Both HopQII strains X47 and SS1 did not show a signal for the antibody against HopQI, which confirmed the specificity of the antibody. Surprisingly, PMSS1 was positive for HopQI in the blot (Figure 17 B). PMSS1 is the original clinical isolate that was passaged through mouse and resulted in the isolation of SS1. After sequencing the strains PMSS1 and SS1, it was confirmed that PMSS1 possessed a functional HopQI in both the *sabA* and *sabB* locus as well as a HopQII in the *hopQ* locus, whereas SS1 produced only a functional HopQII in the *hopQ* locus (supplementary data). The HopQI protein of both the *sabA* and *sabB* locus in the SS1 strain is not synthesized due to phase variation. HopQ is described to be able to switch loci between *hopQ*, *sabA* and *sabB*<sup>126</sup>.

P12 wildtype, P12 $\Delta$ *hopQ* and P12 $\Delta$ *hopQ*::*hopQI* were also examined for HopQI production. The western blot analysis for the HopQI production confirmed the results obtained for the binding to CEACAMs in the bacterial pull-down assays (Figure 15 B). AK298 did not recognize a signal for the P12 $\Delta$ *hopQ* deletion mutant in the western blot. The genetically complemented strain P12 $\Delta$ *hopQ*::*hopQI* was synthesizing HopQI and produced an even stronger signal when compared to the P12 wildtype strain (Figure 17 C) which is consistent with the significantly stronger CEACAM-binding in the bacterial pull-down (Figure 15 C).

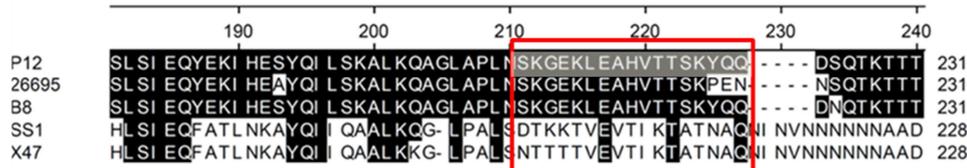
To examine, if this was also true for other strains, B8 was deleted of HopQ and afterwards complemented with the *pHel3* shuttle plasmid pCE39. The strain B8 had two identical *hopQI* gene copies in the loci of *hopQ* and *sabB*. In B8, one *hopQI* gene present in the *hopQ* gene locus

---

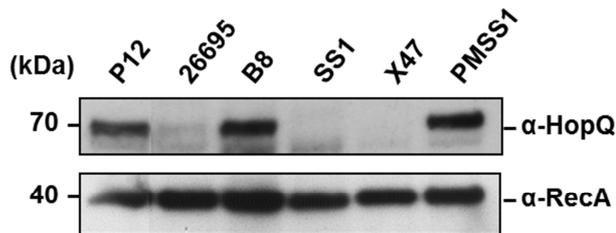
was deleted (*B8ΔhopQ1*). The second copy of *hopQI* still remained in the *sabB* gene locus <sup>144</sup>. *B8ΔhopQ1* was complemented with the plasmid pCE39 as was the *P12ΔhopQ* deletion mutant (Figure 17 D) resulting in a *hopQI* (P12) complemented *B8ΔhopQ1* strain (*B8ΔhopQ1::hopQI* (P12)).

The examination of the HopQI in the western blot revealed for the *B8ΔhopQ1* deletion mutant a weak band at the size of HopQI (Figure 17 C). Since the second gene copy in the *sabB* gene locus is still intact, HopQI was synthesized, but the level was clearly reduced compared to the B8 wildtype level. The *B8ΔhopQ1::hopQI* (P12) strain produced a significantly stronger HopQI signal than the B8 wildtype strain which is consistent with the results obtained for *P12ΔhopQ::hopQI* and P12 wildtype.

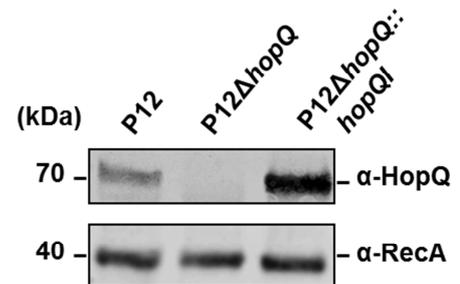
A



B



C



D

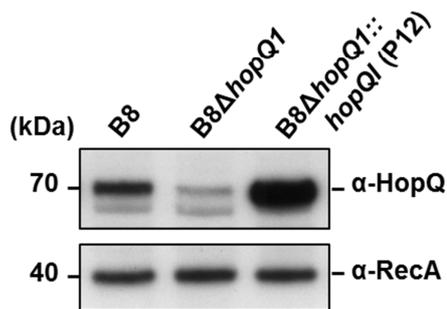


Figure 17: Generation and characterization of the antibody AK298 against HopQI.

(A) An excerpt of the HopQ alignment of strains P12, 26695, B8, SS1 and X47 is depicted. The sequence between aa 211 and 227 of strain P12 (grey) shows the peptide the HopQI antibody AK298 is made against (red box). (B) The strains P12, 26695, B8, SS1, X47 and PMSS1 were tested for HopQI production by western blot analysis. Furthermore, the specificity of the antibody for HopQI was analyzed by examination of HopQI negative strains. The HopQI production is compared to the production of the bacterial protein RecA (40 kDa). (C) The production of HopQI in the P12 wildtype, P12ΔhopQ deletion mutant and the genetically complemented P12ΔhopQ::hopQI is examined. (D) The results for HopQI production of strain B8 complemented with plasmid pCE39 (B8ΔhopQ1::hopQI (P12)) is compared to the B8 wildtype strain and the deletion mutant B8ΔhopQ1.

## 3 CEACAM producing cell lines and CagA translocation

### 3.1 TEST FOR CEACAM PRODUCTION ON CELL LINES

It was shown that *H. pylori* interacts with the human receptors CEACAM1 and CEACAM5. CEACAM5 is exclusively synthesized on epithelial cells, whereas CEACAM1 is found in various cell types like epithelial and immune cells<sup>50</sup>. To study the impact of CEACAMs on the bacteria-host interactions, different cell lines were examined for the production of CEACAMs on the cell surface. The focus was placed on epithelial cell lines, since both CEACAM1 and CEACAM5 were characteristic for this cell type and *H. pylori* is known to target these cells<sup>145</sup>.

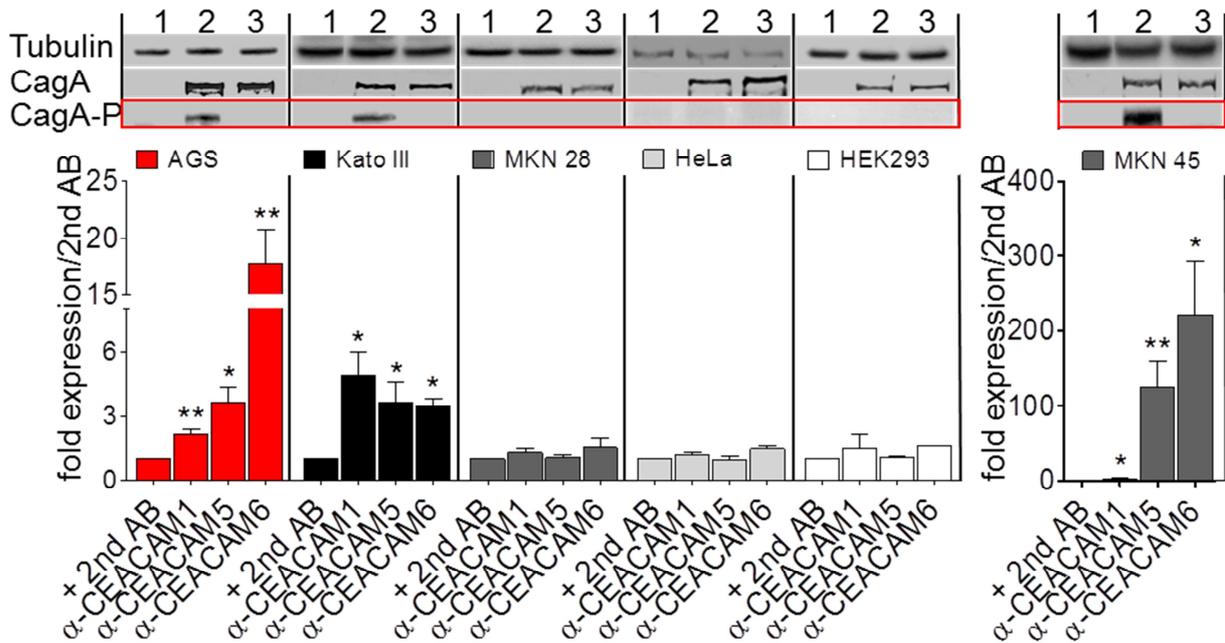
The AGS cell line, a gastric cancer cell line, was first evaluated for CEACAM production. Besides, the gastric cancer cell lines Kato III, MKN28 and MKN45 were tested which represent cell lines originating from tissue being targeted by *H. pylori*. In contrast, the cervix carcinoma cell line HeLa and the embryonic kidney cell line HEK293 were originating from tissue that *H. pylori* is not described to address.

CEACAM1, CEACAM5 and CEACAM6 are typically found on epithelial cells and represent the receptor members known to be targeted by pathogenic bacteria<sup>25</sup>. Moreover, these three CEACAM family members are mostly coproduced on these cells<sup>130</sup>. Although CEACAM6 did not show an interaction with *H. pylori* in the bacterial pull-down assay, the CEACAM6 production was analyzed. The cell lines were stained with specific antibodies against CEACAM1, CEACAM5 and CEACAM6 to analyze the fold production of CEACAMs on the cell surface by flow cytometry and to compare it to the negative control. The flow cytometer readout showed that AGS, Kato III and MKN45 cells were significantly producing CEACAMs. In contrast, MKN28, HeLa and HEK293 cells were negative for the tested CEACAMs (Figure 18).

The study of AGS cells revealed for CEACAM1 2.1 fold and for CEACAM5 3.7 fold the level of the negative control, whereas CEACAM6 was more pronounced on the cell surface with a level of 17.7 fold the negative control.

Kato III cells showed a different pattern than AGS cells. The synthesis level for CEACAM1, CEACAM5 and CEACAM6 was about equally high with mean values of 4.9 fold, 3.6 fold and 3.5 fold compared to the negative control. Thus, Kato III cells exhibited the highest CEACAM1 level of the tested cell lines. CEACAM6 synthesis in Kato III cells was clearly reduced compared to the synthesis in AGS cells (20% of the level in AGS cells).

When having a closer look on MKN45 cells, the similarity of the CEACAM production pattern to AGS cells became apparent, but the values for CEACAM5 and CEACAM6 were significantly higher in MKN45 cells. CEACAM1 was produced with 2.3 fold, CEACAM5 had a production level of 124.9 fold and CEACAM6 of 221.7 fold the negative control.



**Figure 18: Synthesis of CEACAM1, 5 and 6 on epithelial cell lines.**

Six epithelial cell lines standardly used to analyze host-pathogen interactions were examined for the production of CEACAM1, CEACAM5 and CEACAM6. AGS, Kato III, MKN28, MKN45, HeLa and HEK293 cells were stained with specific antibodies against CEACAM1, CEACAM5 and CEACAM6, respectively. The fold production of CEACAMs was detected by flow cytometry as MFI values and compared to the value of the negative control which was a sample incubated only with the secondary antibody and set to the value of 1.

On top, the evaluation of the *in vitro* infection for each cell line is depicted by western blot analysis. Each cell line was infected with *H. pylori* strain P12 (2) and the CagA translocation deficient strain P12Δ*cagI* (3). The mock control consisted of a sample of cells alone (1) that was not infected with bacteria.

Tubulin was used as control that equal amounts of cells were analyzed during the infection with *H. pylori*. AK268 was directed against the N-terminal part of CagA and stained bacterial CagA. Antibody 4G10 showed the level of CagA phosphorylation (red box) e.g. the amount of translocated CagA tyrosine phosphorylated by cellular kinases.

Student's T-test; \* $p \leq 0.05$ , \*\* $p \leq 0.01$ . Values are means +/- SD, n = 3

Since HopQ is the adhesin interacting with CEACAM1 and CEACAM5, respectively, and has recently been described as T4SS-associated virulence factor<sup>140</sup>, it was interesting to look at a possible correlation between *in vitro* infection by *H. pylori* and CEACAM synthesis on the cell surface.

Cells positive and negative for CEACAMs were examined for the capability of *H. pylori* to infect these cells *in vitro*. The success of the *in vitro* infection was determined by the CagA translocation capability into the host cells, e.g. the amount of CagA that was tyrosine phosphorylated by the host cell kinases using the already established method of CagA

phosphorylation assay<sup>108</sup>. The tested cell lines that were capable of CEACAM synthesis were also positive for CagA translocation. The cells that were negative for CEACAM production did not show a phosphorylation of CagA. Figure 18 shows on top the *in vitro* infection test of six different cell lines. Antibody 4G10 shows the level of tyrosine phosphorylated CagA (red box). The six cell lines were infected with strain P12 (2) and the CagA translocation deficient strain P12 $\Delta$ *cagI* (3) that served as negative control, since CagI was described to be necessary for CagA translocation<sup>146</sup>. The negative control consisted of a sample of cells alone (1) that was not infected with bacteria. The strongest CagA translocation was found for MKN45 cells that were also producing the highest amounts of CEACAM5 and CEACAM6 on the cell surface. Kato III cells showed a similar CagA phosphorylation efficiency to AGS cells.

### 3.2 CEACAM PRODUCING HEK293 CELLS AND CAG<sub>A</sub> TRANSLOCATION

Cell lines positive for the synthesis of CEACAM1, CEACAM5 and CEACAM6 correlated with their respective CagA translocation efficiency. Cell lines lacking CEACAMs on the cell surface did not show CagA translocation. To address the question if there is a functional dependency between CEACAM production and CagA translocation, a cell line that was negative for CEACAMs on the cell surface was equipped with a CEACAM member and checked for CagA translocation capability.

HEK293 cells stably transfected with CEACAM1 and CEACAM5, respectively, were obtained from the lab of Jürgen Heesemann (LMU, Munich) <sup>147</sup>. The stably transfected cells were first tested for the production level of CEACAMs by flow cytometry. HEK293/CEACAM1 cells were lacking CEACAM5 or CEACAM6, but were significantly positive for CEACAM1 with a 6.6 fold production compared to the mock control. Nontransfected HEK293 cells were negative for all three tested CEACAMs (Figure 19 A). The evaluation of the flow cytometry data showed for the transfected HEK293/CEACAM1 cells the highest level of CEACAM1 production of the tested cell lines.

Nontransfected control- and HEK293/CEACAM1 cells were infected with *H. pylori* strains P12 (Figure 19 A, lane 2) and the CagA translocation deficient strain P12 $\Delta$ cagI (Figure 19 A, lane 3). The mock control consisted of a sample with cells alone (Figure 19 A, lane 1) that was not infected with bacteria. The CagA phosphorylation assay showed a clear reconstitution of the CagA translocation in HEK293/CEACAM1 cells with a distinct band for the phosphorylation of CagA (Figure 19 A, red box).

For the transfected HEK293/CEACAM5 cells, the flow cytometry readout showed a considerable rise of the CEACAM5 production level to 116.8 fold compared to the negative control, which is about the CEACAM5 level of normal MKN45 cells (Figure 18). No synthesis of CEACAM1 or CEACAM6 was registered in these cells as anticipated (Figure 19 B).

The test for CagA translocation into HEK293/CEACAM5 cells was also positive as it was for HEK293/CEACAM1 cells, but HEK293/CEACAM5 cells had a significantly stronger CagA phosphorylation signal than HEK293/CEACAM1 cells (Figure 19 B, red box).

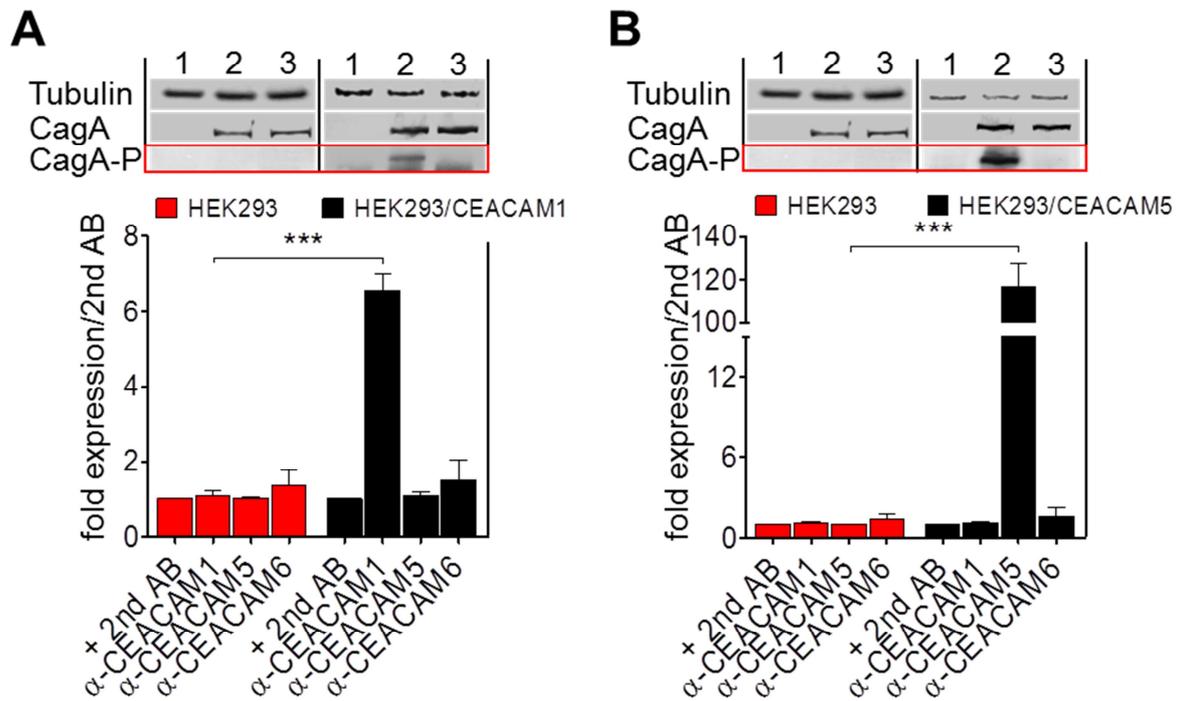


Figure 19: HEK293 cells producing human CEACAM1 and CEACAM5, respectively.

HEK293/CEACAM1 (A) and HEK293/CEACAM5 cells (B) were tested by flow cytometry for the production of CEACAM1, CEACAM5 and CEACAM6 and compared to nontransfected HEK293 control cells.

HEK293/CEACAM1 and HEK293/CEACAM5 cells were infected with *H. pylori* strain P12 (lane 2) and the CagA translocation deficient strain P12 $\Delta$ cagI (lane 3) and compared to the nontransfected control cells. The mock control (lane 1) was a sample of cells alone that was not infected with bacteria. The infection was analyzed using the CagA phosphorylation assay. The translocated CagA was phosphorylated by cellular kinases. Tubulin was used as control for equal amounts of cells. AK268 stained bacterial CagA. Antibody 4G10 showed the level of CagA phosphorylation (red box). Student's T-test; \*\*\* $p \leq 0.001$ . Values are means +/- SD,  $n = 3$

### 3.3 *IN VITRO* INFECTIONS WITH P12 $\Delta$ HOPQ MUTANT AND THE HOPQI AND HOPQII COMPLEMENTED STRAIN

The OMP HopQ is binding to the cellular receptors CEACAM1 and CEACAM5 which are typically produced on epithelial cells. There is also proof of a direct correlation between CEACAM1 and CEACAM5 synthesis on cell lines and CagA translocation capability. HEK293/CEACAM1 and HEK293/CEACAM5 cells showed a reconstituted and strong signal for CagA translocation compared to nontransfected HEK293 control cells. To study the HopQ influence on CagA translocation, HEK293/CEACAM5 and AGS cells were infected with P12 wildtype, the deletion mutant P12 $\Delta$ hopQ and the genetically complemented mutant strain P12 $\Delta$ hopQ::hopQI. A sample of uninfected HEK293/CEACAM5 cells served as negative control. The *in vitro* infections were evaluated by CagA phosphorylation assay and western blot analysis.

HEK293/CEACAM5 cells infected with the P12 wildtype revealed a clear band for CagA phosphorylation, while the signal was completely gone in the P12 $\Delta$ hopQ deletion mutant. The CagA phosphorylation was reconstituted in the strain P12 $\Delta$ hopQ::hopQI. The phosphorylation band between P12 wildtype and P12 $\Delta$ hopQ::hopQI infected cells was about the same strength, although HopQI was significantly stronger produced in P12 $\Delta$ hopQ::hopQI compared to P12 wildtype strain as it was seen in the blot (Figure 20 A). The results suggested for the CagA translocation in HEK293/CEACAM5 cells an absolute dependency on the outer membrane protein HopQ.

Since HEK293/CEACAM5 cells were generated by transfection, the effect was also studied in human gastric AGS cells infected with P12 wildtype, the deletion mutant P12 $\Delta$ hopQ or the genetically complemented mutant strains P12 $\Delta$ hopQ::hopQI and P12 $\Delta$ hopQ::hopQII. The strain P12 $\Delta$ cagl served as negative control. The cells were treated as described for HEK293/CEACAM5 cells in the *in vitro* infection and the results are shown in Figure 20 B and C. In accordance to the HEK293/CEACAM5 cells, the AGS cells showed the same pattern for the P12 wildtype, P12 $\Delta$ hopQ and P12 $\Delta$ hopQ::hopQI in the *in vitro* infection. In contrast to the HEK293/CEACAM5 cells, the phosphorylation signal for CagA in the P12 $\Delta$ hopQ deletion mutant was considerably reduced in AGS cells, but not completely lost. The hopQI complemented strain P12 $\Delta$ hopQ::hopQI had a significantly stronger CagA phosphorylation signal compared to P12 wildtype. This was consistent with the increase of HopQI synthesis observed in the western blot for P12 $\Delta$ hopQ::hopQI. The hopQII complemented strain P12 $\Delta$ hopQ::hopQII was also capable of CagA phosphorylation and the signal was equal in intensity to the P12 wildtype. Taken together, the results suggested that the CagA translocation in AGS cells is only partially dependent upon HopQ.

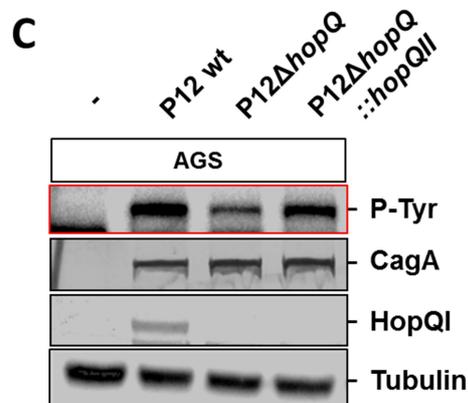
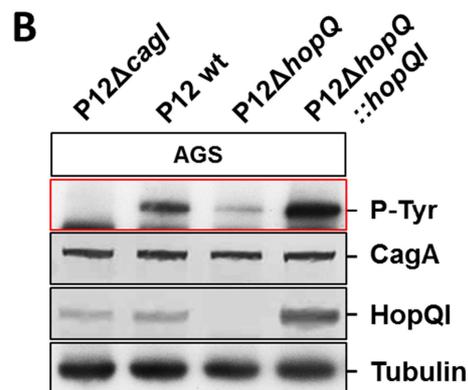
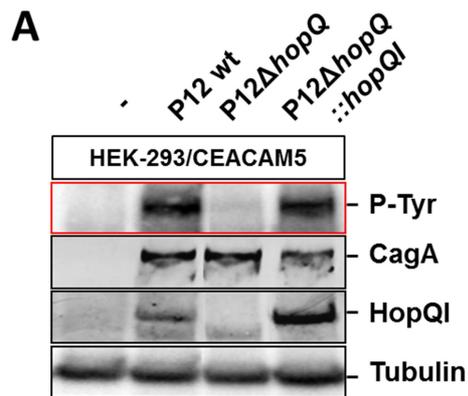


Figure 20: *In vitro* infection of HEK293/CEACAM5 and AGS cells.

(A) HEK293/CEACAM5 cells were infected with P12 wildtype, the P12 $\Delta$ hopQ deletion mutant or the genetically complemented strain P12 $\Delta$ hopQ::hopQI and compared in their capability to translocate CagA into the cells. A sample of uninfected HEK293/CEACAM5 cells served as negative control.

(B) The *in vitro* infection of AGS cells is depicted. P12 wildtype, P12 $\Delta$ hopQ and P12 $\Delta$ hopQ::hopQI were analyzed for CagA translocation efficiency. AGS cells infected with the mutant strain P12 $\Delta$ cagl served as negative control.

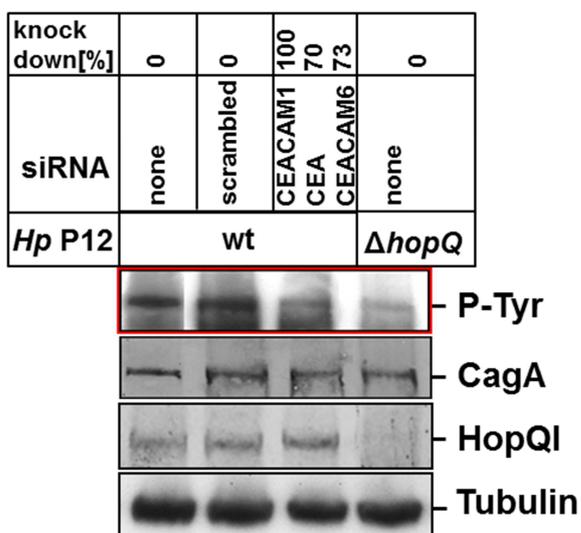
(C) The *in vitro* infection capability of AGS cells for the P12 wildtype, P12 $\Delta$ hopQ and P12 $\Delta$ hopQ::hopQII was compared and evaluated by CagA translocation efficiency. A sample of uninfected AGS cells served as negative control.

Tubulin was used as control for equal amounts of cells. AK268 stained bacterial CagA. Antibody 4G10 showed the level of tyrosine phosphorylated CagA (red box). AK298 recognized HopQI.

### 3.4 KNOCKDOWN OF CEACAM1, 5 AND 6 IN AGS CELLS

Previous experiments showed for AGS cells infected with P12 $\Delta$ hopQ only a significant reduction of the CagA tyrosine phosphorylation, but not a lost signal. In P12 $\Delta$ hopQ::hopQI, the CagA translocation was not only fully restored when compared to the P12 $\Delta$ hopQ deletion mutant, but even stronger than in the P12 wildtype. To further add evidence to the impact of HopQ on the CagA translocation, siRNA against CEACAMs was used to downregulate the receptor molecules in AGS cells. The Allstars Hs Cell Death Control siRNA served as transfection control which targets essential genes for cell survival. The successful siRNA transfection is reflected by dead cells in the wells and with the Allstars Hs Cell Death Control siRNA a transfection efficiency of about 95% was achieved. The Allstars Negative control siRNA was used as negative control for the siRNA transfection consisting of a scrambled sequence that is not complementary to any known mammalian gene and thus, could not downregulate genes.

Hs\_CEACAM1\_11 siRNA was used for the knockdown of CEACAM1, CEACAM5 and CEACAM6, since the siRNA had a knockdown effect on CEACAM1, CEACAM5 and CEACAM6 (Figure 21).



**Figure 21: Knockdown of CEACAM1, CEACAM5 and CEACAM6 in AGS cells.**

Hs\_CEACAM1\_11 siRNA was transfected in AGS cells to downregulate CEACAM1, CEACAM5 and CEACAM6. The CEACAM knockdown AGS cells were analyzed for CagA translocation efficiency in the *in vitro* infection with P12 wildtype. Nontransfected AGS cells, nontransfected AGS cells infected with P12 $\Delta$ hopQ and AGS cells transfected with Allstars Negative control siRNA served as controls.

AGS cells were transfected with the siRNAs for 48 h and then incubated with specific antibodies against CEACAM1, CEACAM5 and CEACAM6. The knockdown effect was analyzed by flow cytometry and the siRNA transfected AGS cells were compared to nontransfected AGS control cells.

The Allstars Negative control siRNA transfected AGS cells did not show a change in the production of CEACAM1, CEACAM5 or CEACAM6. Hs\_CEACAM1\_11 siRNA transfected AGS cells

had a knockdown efficiency of 100% for CEACAM1. For CEACAM5 and CEACAM6, the knockdown effect was prominent, but not complete. A downregulation of 70% for the CEACAM5 and 83% for the CEACAM6 production was achieved.

The CEACAM1, CEACAM5 and CEACAM6 knockdown AGS cells were infected with P12 wildtype and compared to scrambled siRNA transfected and nontransfected AGS cells. The CEACAM knockdown AGS cells showed a clear reduction of the CagA tyrosine phosphorylation signal compared to the control cells. The intensity of the tyrosine phosphorylation in CEACAM knockdown AGS cells matched the band intensity of nontransfected AGS cells infected with the P12 $\Delta$ *hopQ* deletion mutant. This experiment further confirmed that the CagA translocation in AGS cells is not completely dependent on the interaction of CEACAM1/5 and the adhesin HopQ, but additional factors have to contribute to the CagA translocation.

# Discussion

---

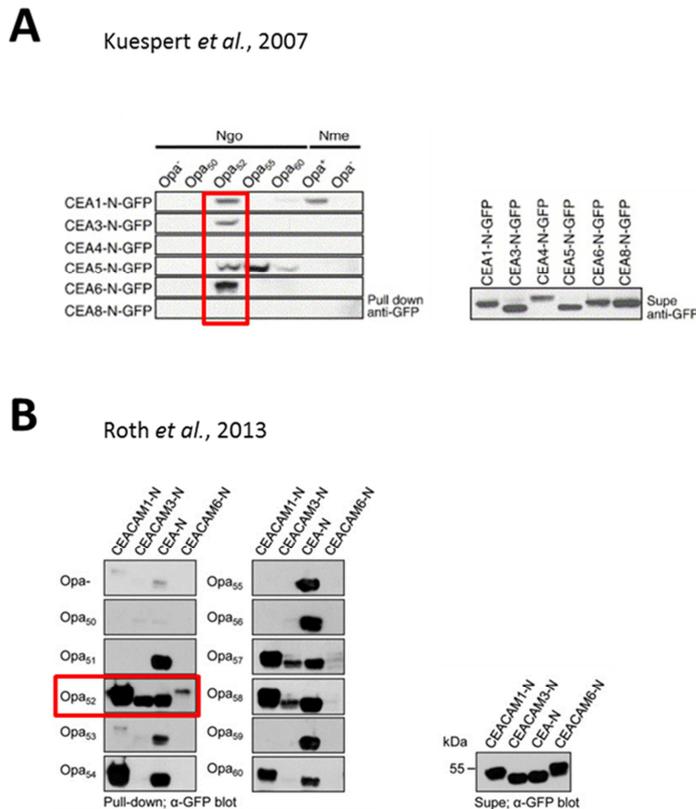
## 1 The capacity of bacteria for CEACAM-binding

### 1.1 THE BACTERIAL PULL-DOWN ASSAY – STRENGTH AND WEAKNESS

Many pathogenic bacteria have been identified to interact with members of the CEACAM family of cell surface receptors. Phase variable opacity-associated (Opa) outer membrane proteins of *N. gonorrhoeae* have been reported to bind to epithelial CEACAM1, CEACAM5 and CEACAM6 and the granulocyte receptor CEACAM3 in CEACAM transfected HeLa cells <sup>148</sup>. In contrast, no interaction occurred between bacteria and CEACAM4, CEACAM7 and CEACAM8, respectively, which indicates a highly specific binding mode <sup>149</sup>. The ubiquitous surface protein A1 (UspA1) adhesin of *M. catarrhalis* showed a similar binding ability regarding CEACAM1, 3, 5 and 6 <sup>78 150</sup>. The group of Christof Hauck applied a bacterial pull-down assay instead of CEACAM transfected cell lines to test for CEACAM-binding. This proved to be a sensitive and sufficient method to test the binding behavior of *N. gonorrhoeae*. The distinct N-terminal IgV-like domain was analyzed for CEACAM-binding. The results were consistent with the already existing data of CEACAM-transfected cell lines <sup>131</sup>. In this work, the bacterial pull-down assay was successfully applied to analyze the capability of *H. pylori* to adhere to CEACAMs. *H. pylori* is, like *N. gonorrhoeae* and *M. catarrhalis*, a Gram-negative and human-restricted, pathogenic bacterium.

The reliability of the bacterial pull-down assay established in this work was tested by published positive and negative controls for CEACAM-binding. The constitutively synthesized Opa 52 protein (Opa<sub>52</sub>) produced by *N. gonorrhoeae* strain MS11 (N309) and the UspA1-producing *M. catarrhalis* strain ATCC 25238 were tested for adherence to CEACAM1, 3, 5 and 6, respectively. Both, *N. gonorrhoeae* strain N302 and *M. catarrhalis* strain ATCC 423517, were not able to bind to any CEACAMs and were used as negative controls <sup>130 43</sup>. As anticipated, the adhesin-negative strains did not show an association with CEACAMs. In contrast, *N. gonorrhoeae* N309 and *M. catarrhalis* strain ATCC 25238 were interacting with CEACAM1, 3 and 5, respectively, however, the binding to CEACAM6 could not be confirmed in this work (Figure 10 B and C). A comparison of the published data of Kuespert *et al.*, 2007 and Roth *et al.*, 2013 revealed discrepancies regarding the intensity of CEACAM6-binding (Figure 22). In Kuespert *et al.*, 2007, the interaction between the adhesin Opa<sub>52</sub> and CEACAM6 was found to be the strongest of all CEACAMs, whilst, in Roth *et al.*, 2013, the binding intensity of Opa<sub>52</sub> has significantly decreased for CEACAM6 to a background level when compared to the binding intensity of

CEACAM1, 3 and 5, respectively. The results obtained in this work for the *Neisserial* binding of CEACAM1, 3, 5 and 6 by Opa<sub>52</sub> are mostly consistent with the recently published data of Roth *et al.*, 2013. Possible explanations for this behavior might be that the CEACAM6 construct tends to misfolding or posttranslational processing.



**Figure 22:** Excerpt from Kuespert *et al.*, 2007 and Roth *et al.*, 2013.

The western blot analysis shows the binding capability of distinct Opa proteins to CEA-N-GFP constructs tested by bacterial pull-down assay. In both experiments equal amounts of constructs were used in the assay (below).

(A) Published in Kuespert *et al.*, 2007: *N. gonorrhoeae* (*Ngo*) strain MS11 constitutively expressing distinct phase variable Opa proteins and *N. meningitidis* (*Nme*) are depicted.

(B) Published in Roth *et al.*, 2013: The Opa proteins constitutively expressed in *N. gonorrhoeae* strain MS11 and the binding to CEA1-, CEA3-, CEA5- and CEA6-N-GFP (CEACAM-N) is shown.

The binding epitope directly involved in *Neisserial* Opa<sub>52</sub> binding has already been identified. The residues 31 – 41 which form the conserved C-strand of the IgV-like domain together with the CC' loop correlate with strong bacterial binding and are crucial for the *Neisserial* interaction with CEACAMs. In addition, it is interesting that an exchange of residues 27 – 29 in CEACAM6 by the corresponding sequences from either CEACAM1 or CEACAM5 resulted in a binding of Opa proteins that formerly had failed to interact with CEACAM6<sup>149</sup>. Thus, these slight changes in the CEACAM6 sequence already had a major influence on the binding properties of *N. gonorrhoeae*.

*N. gonorrhoeae* N309 bound with the highest intensity to CEACAM1, whereas *M. catarrhalis* strain ATCC 25238 was adhering best to CEACAM5 in the bacterial pull-down assay (Figure 10 B and C). The reason for this might be that the adhesins Opa<sub>52</sub> of *N. gonorrhoeae* and UspA1 of

*M. catarrhalis* are phylogenetically completely unrelated and may vary in their binding epitopes recognized on the targeted CEACAM receptors.

Taken together, the bacterial pull-down assay was tested by the CEACAM-associating bacteria *N. gonorrhoeae* and *M. catarrhalis* which recognized the constructs CEA1-N-GFP, CEA3-N-GFP and CEA5-N-GFP. The established assay was used to examine the CEACAM-binding capability in *H. pylori*.

## 1.2 CEACAM-BINDING OF *H. PYLORI* STRAINS AND THE EFFECT OF ANIMAL PASSAGE

The data obtained for *H. pylori* strain P12 by flow cytometry and western blot analysis were consistent and revealed a clear binding capability for CEACAM1 and CEACAM5, but not for any other CEACAM construct (Figure 11 A, B and C). Other *H. pylori* strains (26695, B8, PMSS1, SS1 and X47) tested in the bacterial pull-down assay also attached to CEACAM1 and CEACAM5 (Figure 12 A and C), however, the intensity of CEACAM-binding varied for each strain. The highest binding capacity was observed in PMSS1, whereas the weakest interaction, especially for CEACAM1, was found in strain X47 adapted to the mouse stomach.

In contrast to *N. gonorrhoeae* and *M. catarrhalis*, no *H. pylori* strain was able to significantly bind the granulocyte receptor CEACAM3 or the epithelial receptor CEACAM6. CEACAM3 does not support cell-cell adhesion, but triggers opsonin-independent phagocytosis and subsequent elimination of bacteria, as it was shown for *N. gonorrhoeae*<sup>44</sup>. The missing CEACAM3 recognition of *H. pylori* strains might be the result of an evolutionary new adaptation to the host. In Berger *et al.*, 2004, the pathogen *E. coli* shows an interaction with CEACAM1, 5 and 6, but not with CEACAM3<sup>151</sup>. In avoiding CEACAM3, bacteria might escape the recognition by the innate immune system and therefore reduce the risk of being eliminated. In *N. gonorrhoeae* only strains that produce Opa<sub>52</sub>, Opa<sub>57</sub> and Opa<sub>58</sub> are binding to CEACAM3, other Opa proteins recognize CEACAM1 and/or CEACAM5 only<sup>130</sup>. This suggests that bacterial adhesins have evolved to differentiate between the granulocyte receptor and epithelial CEACAM members.

Data obtained in our working group for CEACAM3-transfected HEK293 cells support the capability of *H. pylori* strains to bind to CEACAM3. The strain was able to translocate CagA as compared to nontransfected HEK293 control cells (Lea Holsten, unpublished data). The lack of CEACAM3 binding observed for *H. pylori* strains in the bacterial pull-down assay performed in this work might be explained by sensitivity limitation, or differences in the structure between the soluble and the membrane associated forms of CEACAM3. Since the CEA3-N-GFP construct showed a binding for *N. gonorrhoeae* and *M. catarrhalis*, this might suggest that only a minor amount of the *H. pylori* population interacts with CEA3-N-GFP, which is not detected by the bacterial pull-down assay.

Opposite to CEACAM3, the failure in recognition of CEACAM6 might be merely false negative. The CEA6-N-GFP construct was not bound by *N. gonorrhoeae* and *M. catarrhalis* which have already been described to do so. This leads to the assumption that most likely the soluble construct is defective. The capability of *H. pylori* strains to interact with CEACAM6 should be analyzed by the established method of CEACAM transfected cell lines<sup>25</sup>.

All examined *H. pylori* strains, except for B8, were consistent in the distinctly stronger binding ability for CEACAM5 as compared to CEACAM1. For P12, the evaluation of three independent bacterial pull-down assays showed a 4.4 fold stronger intensity in interacting with CEACAM5 as compared to CEACAM1 (Figure 11 D). This was consistent with the data found for *M. catarrhalis* which also adhered with higher intensity to CEACAM5, as compared to CEACAM1. Since the CEACAM1 and 5 binding pattern for *N. gonorrhoeae* was vice versa, the observed preference in adhering to CEACAM5 suggests an adaptation of the *Neisserial* adhesins to the CEACAM5 binding epitope with a possible direct advantage for the bacteria.

In *N. gonorrhoeae* and *M. catarrhalis*, the CEACAM-binding adhesins are phase variable and, therefore, the bacterial population is in general heterogeneous in its ability to recognize CEACAMs<sup>152 153</sup>. The flow cytometry results for the tested *H. pylori* strains showed that the whole bacterial population interacts with CEACAM5 (Figure 12 B). This would suggest that the adhesin(s) responsible for CEACAM5 interaction in *H. pylori* are permanently produced and that the bacterial population is homogeneous regarding the adhesin(s), which is in contrast to *N. gonorrhoeae* and *M. catarrhalis*.

Regarding CEACAM1, a homogeneous population was observed for the strains P12 and 26695. The strains SS1 and X47 adapted to the mouse model were significantly impaired regarding the association with CEACAM1. Only about 10% of the P12 binding intensity was detected when adhering to CEACAM1 (Figure 12 B). This suggests a reduced adhesin(s) production or bacteria that express adhesin(s) with a weaker CEACAM1-binding capability. The allele for the murine CEACAM1 orthologue was duplicated resulting in the two proteins CEACAM1<sup>a</sup> and CEACAM1<sup>b</sup> which differ significantly in aa sequence from the human CEACAM1 (45% and 55% aa sequence identity, respectively)<sup>79</sup>. Thus, an interaction between *H. pylori* and the murine CEACAM1 was not likely leading to a reduced HopQ production due to missing selective pressure. CEACAM5 orthologues have not been identified in rodents, but nevertheless the CEACAM5-binding intensity did not change in any *H. pylori* strain.

To analyze the effect of mouse passage on the distinct CEACAM-interaction capability in more detail, strain PMSS1 was examined and compared to strain SS1. The *H. pylori* strain PMSS1 had been passaged through mouse resulting in the strain SS1<sup>132</sup>. In the bacterial pull-down assay, PMSS1 and SS1 were both able to adhere to CEACAM1 and CEACAM5, respectively, but to varying degrees. SS1 had a three fold decreased CEACAM-binding intensity when compared to the non-mouse-passaged strain PMSS1 (Figure 12 C). This means that after animal passage the bacteria had lost their CEACAM-binding ability to a large extent. The CEACAM-binding intensity obtained for X47, another strain adapted to the mouse model, is even lower than that of SS1. This suggests that *H. pylori* strains may lose their CEACAM-binding capability by mouse passage.

The reason for this might be low selective pressure and the downregulation of the corresponding adhesin(s) due to a failure of adherence to murine CEACAM1.

In the gerbil adapted strain B8, the CEACAM1-binding intensity was decreased to about 80% of the P12 strain. When it is assumed that the adhesin(s) responsible for CEACAM1-binding ability is impaired in strains after animal passage, the adhesin(s) in B8 is only slightly reduced after gerbil passage as compared to the mouse passaged strains. CEACAM1 and CEACAM5 orthologues in gerbils have not been addressed yet. Since CEACAM1 has been predicted to be the primordial CEACAM member from which a diverse group of other CEACAMs was generated by gene conversion and duplication<sup>76</sup>, a CEACAM1 protein is most likely also present in gerbils. The comparison of strains PMSS1 and SS1 reveals that, though the bacterial population is homogeneous, the intensity of CEACAM5-binding is significantly lowered after animal passage to 30%, as it was occurring for CEACAM1-binding.

Taken together, these findings suggest regarding the colonization of *H. pylori* strains in mice, a loss of CEACAM-binding ability due to a possible failure in interaction with the CEACAM orthologues in the animals.

### 1.3 CEACAM-BINDING SPECIFICITY

*H. pylori* is able to adhere to human CEACAM1 and 5 in the bacterial pull-down assay. It has to be considered that there is a correlation between the bacterial capability to adhere to CEACAMs and the long term persistence of *H. pylori* in the host. Therefore, three requirements are postulated: (I) establishment of colonization, (II) evasion of the host immunity and (III) invasion of the gastric mucosa<sup>116</sup>. The exploitation of CEACAMs might be favorable, since the receptor family is involved in major cellular processes, like cell-cell adhesion and regulation of the immune system<sup>27</sup>. *H. pylori* strains can colonize animals like gerbils, rhesus monkeys or mice and theoretically may also be capable to exploit CEACAM members present in these animals to establish a bacterial infection in the host. CEACAM1 orthologues have been identified in all analyzed species, whereas CEACAM5 is exclusively present in primates<sup>76</sup>. Recent studies showed for the human-restricted pathogens *N. gonorrhoeae* and *M. catarrhalis* that they do not recognize CEACAM1 orthologues from mouse (CEACAM1<sup>a</sup>), cattle (CEACAM1<sup>a</sup> and CEACAM1<sup>b</sup>) and dog<sup>43</sup>. Since the CEACAM-binding ability after the animal passage was clearly decreased in strain SS1 as compared to PMSS1 (Figure 12 C), the assumption was made that *H. pylori* strains might not be able to interact with orthologous CEACAMs in the murine stomach.

In this work, the examined *H. pylori* strains failed to interact with CEACAM1 and CEACAM5 of other species than humans (Figure 13 B). Although SS1 and X47 were able to colonize mice, the strains do not seem to exploit the murine CEACAM1 receptor, but probably use other adhesion factors. Since there is no positive control existing for the orthologous CEACAM1 and CEACAM5 constructs which would confirm the functionality of the constructs, the incapability in binding is not absolutely confirmed. Moreover, it has to be considered that the CEACAM1 IgV-like domain in different species is not highly conserved. Due to a lacking selection pressure, the domains differ largely. Compared to the human CEACAM1, the murine variant shows an aa sequence conservation of 45%<sup>79</sup>, the bovine variant of 54%<sup>80</sup> and the canine variant of 56%<sup>75</sup>. This might explain the specific restriction of *H. pylori* to its host. Such CEACAM-dependent host specificity has also been found for the murine hepatitis virus. Data obtained from the murine hepatitis virus-infection in mice confirmed the exclusive and specific binding of the virus to the murine CEACAM1<sup>a</sup> IgV-like domain with a binding site that has not been found in the CEACAM1 sequence of any other species<sup>80</sup>.

Non-*pylori Helicobacter* species are also found to colonize animals like rhesus monkeys. In many cases, the bacteria can colonize the gastric epithelium of animals and cause an inflammatory response resembling that seen with *H. pylori* in humans<sup>154</sup>. Despite this similarity in inflammatory response, the non-*pylori* species, *H. nemestrinae* and *H. mustelae*, were not able to establish a binding to the human receptor in the bacterial pull-down assay (Figure 13 A). This

suggests that the binding capability to CEACAMs is a specific adaptation of the species *H. pylori* to its human host.

Taken together, exclusively *H. pylori* strains are able to adhere to the human CEACAM receptor family. Moreover, *H. pylori* is species-specific in the recognition of CEACAMs and is unable to adhere to CEACAMs found in other species than humans.

## 2 The CEACAM-binding adhesin HopQ

### 2.1 HOPQ IS THE ADHESIN INTERACTING WITH CEACAM1 AND CEACAM5

BabA is the bacterial adhesin binding to fucosylated Lewis b antigens and SabA adheres in inflamed tissue to sialyl-Lewis x and a antigens. Both belong to the adhesin branch of the Hop family of OMPs. The adhesin(s) interacting with CEACAM1 and CEACAM5 were assumed to also belong to this family (Figure 14 C), since several proteins in this branch have not been assigned a function yet. Distinct single OMP gene deletion mutants of *H. pylori* strain P12 were generated by the streptomycin contraselection system and analyzed in the bacterial pull-down assay for a loss or reduction in CEACAM-binding compared to the P12 wildtype. The generated deletion mutants for BabA, BabB, SabB and HopZ were not involved in the CEACAM-binding ability, since each of the mutants recognized both CEACAM1 and CEACAM5 with a binding intensity most similar to the one observed in P12 wildtype. In contrast, deletion mutants generated for HopQ in different strain backgrounds completely failed to interact with CEACAM1 and CEACAM5. The complete loss of binding to CEACAMs identified HopQ as the sole adhesin interacting with CEACAM1 and CEACAM5, respectively. When the *hopQ* deletion mutant was genetically complemented with the *hopQ* gene, the adherence to CEACAM1 and CEACAM5 was fully reconstituted.

It was reported that *H. pylori* strains vary in their virulence outcome depending on the allelic form of the *hopQ* gene<sup>105</sup>. *H. pylori* type II strains with a HopQII sequence lack the CagPAI and are less pathogenic than HopQI producing, CagPAI positive type I strains. When comparing *hopQI* and *hopQII* complemented strains, strain P12 $\Delta$ *hopQ*::*hopQII* was significantly weaker in binding to CEACAM1 and CEACAM5 compared to strain P12 $\Delta$ *hopQ*::*hopQI* (Figure 15 D). Type I strains of *H. pylori* seem to adhere stronger to CEACAMs due to the HopQI aa sequence.

The ClustalW alignment of the HopQ aa sequence showed for P12, 26695 and B8 a HopQI and for the mouse adapted strains SS1 and X47 a HopQII sequence (Figure 16 A). The HopQI and HopQII sequences revealed a strong conservation of 50 aa in the N-terminal region and of 200 aa in the C-terminal region of the aa sequence leaving a variable middle region to distinguish between HopQI and HopQII. The HopQI sequences of P12 and 26695 were highly conserved with 96.6% identity, although 26695 showed less binding capability to CEACAMs compared to P12 in the bacterial pull-down assay (Figure 12 A). The HopQI sequence of the gerbil adapted strain B8 differed from the one found in P12 merely in the region between the aa 429 and 478. Interestingly, exactly this region was almost identical to the HopQII sequence found in the mouse adapted strains SS1 and X47 and, therefore, might be advantageous for the colonization in animals. The HopQII sequences of SS1 and X47 differed from the P12 sequence, with an aa identity of about 72% (Figure 16 B), but differed from each other in merely 21 aa (aa sequence

identity of 96.7%). The SS1 capability to bind to CEACAM1 and X47 adhering to CEACAM1 and 5 was significantly reduced compared to the binding found for P12 (Figure 12 A). X47 had the same CEACAM-interaction capability like P12 $\Delta$ *hopQ::hopQII*. This phenotype observed in the tested strains together with the findings for the genetically complemented strains P12 $\Delta$ *hopQ::hopQI* and P12 $\Delta$ *hopQ::hopQII* suggest that, dependent on the HopQ allelic form expressed in the strain, the CEACAM-recognition capability is significantly varying. The strains SS1, X47 and P12 $\Delta$ *hopQ::hopQII* share a HopQII aa sequence and show a reduced CEACAM1-binding intensity compared to the HopQI producing strains P12, 26695 and B8. The CEACAM5-binding intensity does not seem to be influenced by the allelic form of HopQ to the same degree than the CEACAM1-binding capability.

Taken together, both forms of the adhesin HopQ are able to interact with CEACAM1 and CEACAM5, but the HopQII form seems to recognize it weaker. The specific interaction represents a further way of *H. pylori* to adhere to different cell types and tissues of its human host.

## 2.2 MODE OF HOPQ BINDING TO CEACAM5

The adhesins BabA and SabA are reported to bind to carbohydrate structures present on their receptors <sup>124</sup>. In this work, the question has not been addressed whether HopQ binds to carbohydrate structures on the CEACAMs or to the protein backbone of the glycosylated receptors. The CEACAM family members are highly glycosylated proteins with a possible carbohydrate content of up to 60 – 70% in weight, which makes glycosylation a major factor in these proteins <sup>1</sup>. The AGFCC'C''- $\beta$  sheet is predicted to be devoid of carbohydrate structures, whereas the ABDE face is most likely covered by them <sup>155</sup>. In *N. gonorrhoeae* and *M. catarrhalis*, the binding site for the Opa proteins and UspA1, respectively, was identified to be the AGFCC'C''- $\beta$  sheet of CEACAMs and not carbohydrate structures <sup>149 152</sup>.

The respective HopQ binding site on CEACAMs was examined by isothermal titration calorimetry (ITC) (performed in collaboration with the lab of Eric Sundberg, University of Maryland, Baltimore) (unpublished data). The N-terminal domain of CEACAM5 (CEA-N) from *E. coli* was purified, resulting in a CEA-N domain devoid of glycosylation (MW = 11.88 kDa). In addition to that, HopQ was produced as fusion protein with maltose binding protein (MBP) and purified. The non-glycosylated CEA-N was titrated against MBP-HopQ. CEA-N titrated against MBP alone served as negative control. MBP-HopQ showed a high affinity binding for the non-glycosylated CEA-N ( $K_D = 61 \pm 3$  nM) with a stoichiometry of one MBP-HopQ to half of a CEA-N dimer (1:0.62), similar to other adhesin-CEA interactions <sup>156</sup>. MBP alone showed no affinity for the non-glycosylated CEA-N. These findings confirm the high affinity binding observed for CEACAM5 and HopQ in the bacterial pull-down assay. Furthermore, these data show for the first time a direct binding of *H. pylori* to the receptor protein and not to any glycosylated structure found on the receptor. This makes the adhesion of HopQ to CEACAM5 the first protein-protein binding interaction found for *H. pylori*.

It is speculated that the CEACAM-binding site for HopQ is most likely the AGFCC'C'' face of the IgV-like domain (Figure 1) which has already been identified as the binding site for *N. gonorrhoeae* and *M. catarrhalis*. The  $\beta$ -sheet is predicted to be devoid of carbohydrate structures which would match the result of HopQ binding to the protein backbone of CEACAM5. The AGFCC'C'' face of CEACAM1 alone was shown to be less conserved between species as compared to the complete IgV-like domain. Compared to the human variant, the aa sequence conservation of the AGFCC'C'' face of murine and canine CEACAM1 is less than 40% and the bovine one is less than 50% <sup>157</sup>. This would strengthen the hypothesis that the observed lack of interaction between *H. pylori* strains and the orthologous CEACAM1 constructs is due to the high sequence variations in the binding domain.

## 2.3 HOPQ PRODUCTION IN *H. PYLORI* STRAINS

In addition to the allelic type of HopQ, also the amount of HopQ produced in *H. pylori* strains might differ, which can further explain the observed differences in CEACAM-binding. By the generation of a HopQI-specific antibody, the protein production was analyzed in the strains. HopQI was detected in the strains P12 and B8 at relatively equal levels, but weak for strain 26695, whereas the HopQII of strains SS1 and X47 is not recognized by the antibody due to epitope variation (Figure 17 A). When the HopQI sequence of strain 26695 was examined for the specific epitope against which the antibody was generated, it was evident that the last three aa of the peptide sequence were divergent and, therefore, the HopQI protein of strain 26695 might not be fully recognized by the HopQI antibody (Figure 17 B).

In contrast, PMSS1 showing a strong signal for HopQI was discovered to have a duplication in the *hopQI* gene. PMSS1 is the original clinical isolate of SS1 that has been passaged through mouse. When the *hopQ* sequence in the *hopQ* locus of PMSS1 and SS1, the sequences were identical and showed for both a functional *hopQII* gene. Due to the positive western blot result of PMSS1 for HopQI (Figure 17 B), the strain was examined again and the *hopQI* gene sequence was identified both in the *sabA* and *sabB* locus. Interestingly, the *hopQ* gene as well as the closely related *sabA* and *sabB* genes are variable in copy number and locus<sup>126</sup> which can be switched in their loci due to gene conversion. A *hopQI* gene was also identified in strain SS1 in the *sabA* and *sabB* gene loci, but the genes are not expressed. The *hopQI* genes are phase variable due to SSM and HopQI is not detected by western blot analysis. This means that the strain SS1 has lost the HopQI production after it has been passaged and isolated from the mouse indicating that the colonization of mice by this strain was not dependent on the production of HopQI. This matches the observations from the bacterial pull-down assay. The depletion of the HopQI production in SS1 resulted in a CEACAM-binding ability of only 30% of the PMSS1 level which produces a functional HopQI as well as a functional HopQII and showed a significantly stronger CEACAM1 and CEACAM5 interaction (Figure 12 C).

In Talarico *et al.*, 2012, 51 clinical strains were examined for gene conversion, which involved the duplication of the *sabA* or *hopQ* gene and the loss of the *sabB* gene (Figure 24). Eight clinical isolates were positive for the *hopQ* gene duplication. Six comprised one *hopQI* and one *hopQII* gene, whereas two isolates had two copies of a *hopQI* gene that appeared to be the result of a recent intra-strain gene conversion event<sup>126</sup>. It was found that by *sabA* duplication, SabA protein production was two-fold higher and resulted in an increase in adherence to immobilized sialyl-Lewis x. This gain of protein production and the resulting enhancement of receptor binding occurred in strain PMSS1. The duplication of the *hopQI* gene additional to the *hopQII* gene led to a stronger interaction with CEACAMs and from the analyzed strains in this work, PMSS1 had by

far the highest adherence to CEACAMs. It was most interesting to observe that after animal passage the reisolated strain SS1 had lost its HopQI production and the CEACAM-binding intensity was directly correlating with the produced HopQ type. By the loss of HopQI production, only one third of the CEACAM-binding ability remained in strain SS1 compared to the original isolate PMSS1.

A second strain besides PMSS1 showed a duplication of the *hopQI* gene: PMSS1 comprises two identical *hopQI* alleles in the *sabA* and the *sabB* gene locus, whereas B8 contains two identical *hopQI* genes in the *sabB* and *hopQ* gene locus<sup>158</sup>. Although two copies of HopQI are produced in the B8 wildtype and merely one HopQI is produced in P12, the amount of the HopQI protein of B8 and P12 wildtype in the western blot were identical (Figure 17 B). However, the CEACAM-binding capability is lower for the B8 wildtype (Figure 12 A). This difference in CEACAM adherence is most likely explained by variations in the HopQ aa sequence between P12 and B8. The region responsible for the CEACAM-binding might be located between the aa 429 and 477, since in this region, P12 and 26695 are completely identical in sequence. However, B8 and all HopQII producing strains, which show a reduced CEACAM-binding capability, differ from the P12 sequence in this region, but are nearly identical between each other. Further strains and HopQ sequences have to be analyzed to verify this correlation.

The *hopQI* gene located in the *hopQ* locus was deleted in strain B8 by homologous recombination. The HopQI production loss resulting from the *hopQI* gene deletion led to a reduced HopQI production of about 30% in strain B8 $\Delta$ *hopQ1* compared to the B8 wildtype (Figure 17 D). Therefore, the deleted *hopQI* gene was responsible for the main HopQI production in the B8 wildtype. The minor *hopQI* gene expression in the *sabB* locus is possibly explained by a weaker activity of the promoter in the *sabB* locus as compared to the promoter in the *hopQ* locus.

Furthermore, the P12 wildtype, the deletion mutant P12 $\Delta$ *hopQ* and the *hopQI* complemented strain P12 $\Delta$ *hopQ::hopQI* were analyzed for HopQI production. The deletion mutant did not produce a HopQI which matches the loss of CEACAM-binding. In contrast, P12 $\Delta$ *hopQ::hopQI* was significantly overproducing HopQI compared to the production in the P12 wildtype (Figure 12 C). Interestingly, the amount of HopQI in the distinct P12 strain correlated directly with the CEACAM-binding capability in the bacterial pull-down assay (Figure 15 C). The strong increase in HopQI production in P12 $\Delta$ *hopQ::hopQI* was explained by the promoter in front of the *hopQ* gene on the complementation plasmid pCE39. When strain B8 $\Delta$ *hopQ1* was transformed with the same plasmid, the newly generated strain B8 $\Delta$ *hopQ1::hopQI* (P12) showed the same increase in HopQI production (Figure 17 D). The amount of HopQII produced in P12 $\Delta$ *hopQ::hopQII* was not recognized by the  $\alpha$ -HopQI antibody due to its specificity for HopQI, but might also be elevated compared to the respective wildtype. Since the plasmids for the complementation of HopQI and

HopQII are identical in the backbone except for the type of *hopQ* gene that was inserted, the plasmids have most likely the same properties in gene expression.

Taken together, gene conversion events regarding the *hopQ* gene can lead in different strains to a gain (B8) or a loss of function (PMSS1 to SS1), concerning the adhesion properties of *H. pylori*.

## 3 Impact of the bacterial CEACAM interaction on CagA translocation

### 3.1 CEACAM PRODUCTION AND CAG A TRANSLOCATION CAPABILITY

In this work, an interaction between *H. pylori* strains and human CEACAM1 and CEACAM5 has been discovered in the bacterial pull-down assay. Members of the CEACAM family were found on various cell types with CEACAM1 being the most wide-spread member amongst them <sup>27</sup>. The expression of CEACAM1, CEACAM5 and CEACAM6 was found to be upregulated in cancerous tissue, for example in colon and gastric cancer <sup>159</sup>. In order to gain access to host cells, pathogenic bacteria were found to adhere to CEACAM1, 5 and 6 and the granulocyte receptor CEACAM3 which is closely related to the IgV-like domain of CEACAM1 and involved in bacterial uptake and elimination <sup>50</sup>. *H. pylori* persistently colonizes the human gastric epithelium and is *in vitro* capable of adhering to CEACAM1, 5 and 6 which are known to be produced on epithelial cells in the gastrointestinal tract <sup>53</sup>.

To gain more knowledge about the interaction between bacteria and CEACAMs, cultured epithelial cell lines were examined for their amount of CEACAMs produced on the cell surface. Moreover, the cell lines were analyzed for the potential of *H. pylori* to infect them. *H. pylori* adheres to cells and translocates the effector protein CagA into the host cell cytoplasm, where the protein is turned into its active form by tyrosine phosphorylation executed by cellular kinases <sup>145</sup>. In an earlier work, HopQ was found to be essential for CagA translocation and for CagA-mediated host cell responses, such as formation of the *hummingbird* phenotype and cell scattering, but the adhesin did not affect motility or the quantity of bacteria attached to the host cells <sup>140</sup>.

The success of the bacterial infection was determined in this work by the amount of phosphorylated CagA. Three gastric cancer cell lines AGS, Kato III and MKN45, showed a significant production of CEACAM1, 5 and 6, whereas the gastric cancer cell line MKN28, the kidney epithelial cell line HEK293 and the cervix carcinoma cell line HeLa were not synthesizing CEACAMs on the cell surface (Figure 18). Except for Kato III cells, the CEACAM1 level found in the cell lines was lower compared to the level of CEACAM5 and CEACAM6. Furthermore, AGS, Kato III and MKN45 cells were positive for CagA translocation (Figure 18, red box). MKN45 cells were particularly outstanding; MKN45 cells had the highest production of CEACAM5 and 6 on the cell surface and the strongest CagA phosphorylation signal in *in vitro* infections. When the findings for the CEACAM production were compared to the data obtained for the bacterial pull-down assay, it was evident that the bacteria adhered better to CEACAM5 as compared to

CEACAM1 (Figure 12 C) and, furthermore, CEACAM5 was produced at higher or in equal amounts compared to CEACAM1 in all examined epithelial cell lines (Figure 18). This might indicate that bacteria are better adapted to CEACAM5 or prefer the binding to CEACAM5. The reason for this might concern the CagA translocation efficiency.

To determine whether the CagA translocation capability of *H. pylori* was dependent on the CEACAM synthesis, CEACAM negative HEK293 cells were transfected with human CEACAM1 and 5, respectively. The *in vitro* infection of *H. pylori* strain P12 in the transfected HEK293/CEACAM1 and HEK293/CEACAM5 cells revealed a clearly reconstituted CagA phosphorylation signal compared to the nontransfected control cells. This indicates that the CagA translocation is dependent on the interaction of *H. pylori* with CEACAMs. *H. pylori* might exploit the cellular receptors to regulate cellular functions most important for the survival in the gastric epithelium. When HEK293/CEACAM1 cells (Figure 19 A) were compared to HEK293/CEACAM5 cells (Figure 19 B), the CagA phosphorylation intensity was considerably stronger in the CEACAM5 transfected cells. HEK293/CEACAM1 cells might synthesize significantly lesser amounts of CEACAM1 than HEK293/CEACAM5 cells produce CEACAM5, although the cells transfected were originating from the same HEK293 cells (143). The stronger CEACAM5 synthesis might contribute to the pronounced CagA phosphorylation in HEK293/CEACAM5 cells. When the ratio of the CEACAM1 production level was compared to the CEACAM5 production level, CEACAM1 might be the one of the two receptor members that affects the CagA phosphorylation signal the most.

### 3.2 HOPQ DEPENDENCY OF CAGA TRANSLOCATION

BabA has been identified as the adhesin responsible for the adherence of *H. pylori* to Lewis b antigens in gastric epithelium. In inflamed tissue, the synthesis of sialyl-Lewis x and a antigens on the cell surface is increased and these antigens are selectively recognized by the bacterial adhesin SabA<sup>116</sup>. The binding of adhesins to the stomach epithelium reduces the constant danger of bacterial clearance by the stomach peristaltic. Furthermore, the BabA interaction is also beneficial in triggering the production of proinflammatory cytokines and precancer-related factors in a T4SS-dependent way<sup>120</sup>. The binding of adhesins might contribute to important signaling pathways in host cells in addition to bacterial adherence as such.

In this work, HopQ was identified as the adhesin interacting with the CEACAM family. Both CEACAM1 and 5 contribute directly to the capability of the bacteria to translocate CagA into epithelial cells *in vitro*. HEK293/CEACAM5 cells showed a complete HopQ dependency for CagA translocation (Figure 20 A). The data obtained for the gastric cancer cell line AGS were consistent with the findings for HEK293/CEACAM5 cells, although the CagA phosphorylation in AGS cells is not fully HopQ dependent. In infected AGS cells, the strain P12 $\Delta$ hopQ had not completely lost the capability of CagA translocation, but showed a reduced signal for CagA phosphorylation compared to the P12 wildtype (Figure 20 B). The remaining CagA phosphorylation might be due to a further still unidentified interaction of a second bacterial adhesin with another cellular receptor. The results obtained for the knockdown experiment in AGS cells strengthen this hypothesis. The downregulation of the epithelial CEACAMs 1, 5 and 6 in AGS cells led to a reduced CagA phosphorylation signal identical to the level of nontransfected AGS cells infected with strain P12 $\Delta$ hopQ (Figure 21).

Since no signal was detected in HEK293/CEACAM5 cells infected with P12 $\Delta$ hopQ, HEK293/CEACAM5 cells might lack the additional cellular receptor. The *hopQI* complemented strain P12 $\Delta$ hopQ::*hopQI* showed a significant increase in HopQI production compared to the P12 wildtype and consequently produced a significantly stronger CagA phosphorylation signal. Due to two allelic forms of *hopQ* associated with more or less virulent *H. pylori* strains, P12 $\Delta$ hopQ::*hopQII* was also examined in AGS cells. The exact amount of HopQII production in P12 $\Delta$ hopQ::*hopQII* could not be identified, since no antibody was available recognizing HopQII. Nevertheless, AGS cells that were infected with the *hopQII* complemented strain showed a significant CagA phosphorylation signal. When the intensity of the signals in AGS cells was compared, P12 $\Delta$ hopQ::*hopQII* had a weaker CagA phosphorylation than the P12 wildtype (Figure 20 C). In accordance with that, the CEACAM-binding capability obtained for P12 $\Delta$ hopQ::*hopQII* in the bacterial pull-down assay was considerably weaker than the P12 wildtype (Figure 15 D). The CEACAM5 interaction was about 50% reduced, whereas the CEACAM1-binding of

P12 $\Delta$ *hopQ::hopQII* was only 10% the level of the P12 wildtype. The significant drop in the CEACAM1-interaction capability of P12 $\Delta$ *hopQ::hopQII* was not reflected in the CagA phosphorylation signal of the AGS infection. This might be due to the remaining CEACAM5 binding capability of 50%. The HopQII production in *H. pylori* strain P12 $\Delta$ *hopQ::hopQII* is most likely influencing the CEACAM-binding to a higher degree than the CagA phosphorylation capability in *in vitro* infections. In contrast, the higher HopQI production in *H. pylori* strain P12 $\Delta$ *hopQ::hopQI* matches the higher fold binding to CEACAMs in the bacterial pull-down assay (Figure 15 C) as well as the elevated CagA phosphorylation signal in *in vitro* infections (Figure 20 B).

In Belogolova *et al.*, 2013, the P12 $\Delta$ *hopQ* deletion mutant did not reveal a CagA phosphorylation signal in AGS cells. The results for the P12 $\Delta$ *hopQ* strains complemented with *hopQI* and *hopQII*, respectively, were consistent with the findings in this work. Both complemented strains were capable to trigger a CagA phosphorylation signal in infected AGS cells <sup>140</sup>.

Experiments which address the question how the HopQ production in bacteria influences the phosphorylation signal of CagA in host cells have not been performed yet. The lack of HopQ synthesis might impair the translocation of the effector protein CagA into host cells, but it is also possible that the adhesin influences the processing of the effector molecule inside the host cells.

## 4 Outlook

This thesis concentrated on the *H. pylori* adhesin HopQ and the CEACAM protein family acting as receptor for *H. pylori*. The impact of this interaction on the CagA translocation *in vitro* suggested a major role for HopQ as bacterial virulence factor. In the bacterial pull-down assay, the interaction of *H. pylori* with CEACAM1 and CEACAM5, respectively, has been confirmed. Regarding CEACAM6, no interaction has been detected, but the results might be false negative due to a non-functional construct. It is necessary to reassess with another experimental model the interaction between *H. pylori* and CEACAM3 and CEACAM6. Furthermore, the N-terminal binding domain of these CEACAMs is closely related to the N-domain of CEACAM1 and CEACAM5 and the CEACAM-binding bacteria *N. gonorrhoeae* and *M. catarrhalis* are already known to bind to the four CEACAM members.

Since the CagA phosphorylation signal was found to be at least partly dependent on the HopQ binding in *in vitro* experiments of AGS cells, an interesting question to address is why there is a remaining CagA phosphorylation signal in the AGS cell infection with the P12 *hopQ* deletion mutant. To date, several proteins of the putative adhesin branch of the *hop* gene family have not been assigned a function yet. The remaining CagA phosphorylation signal observed in the AGS cell infection might result from the interaction of a second adhesin with a yet unknown cellular receptor. The identification of the origin of the remaining signal may also help to discover why HEK293/CEACAM5 cells show a complete HopQ dependency for CagA translocation in contrast to AGS cells. Furthermore, how the HopQ-binding exactly influences the phosphorylation of CagA, is not known yet. HopQ might impair the translocation of the effector protein CagA into cells, but the adhesin interaction might also influence the processing/phosphorylation of the effector molecule inside host cells. Besides the processing of CagA, questions concerning the impact of HopQ on the inflammatory response of the host cells have not been addressed yet. The adherence of HopQ to the host cells may also trigger changes in the bacterial signaling pathways. The influence of HopQ on other bacterial virulence factors besides CagA, e.g. VacA, has not been tested and might also represent an interesting topic.

As identified in this work, HopQI and HopQII proteins behaved different. The CEACAM-binding capability was weaker for strains carrying the *hopQII* alleles and there were also differences regarding the reduced CagA phosphorylation signal in *in vitro* infection experiments. The *hopQII* gene is typically expressed in *H. pylori* type II strains which are reported as less pathogenic strains in *in vivo* studies. This gives rise to questions about the role of HopQ *in vivo* in general and about the effects the two allelic versions have in animal studies. The obtained data for strains PMSS1 and SS1 have already shown interesting details about the HopQ status. After the animal passage, a loss of HopQI production was registered in strain SS1 compared to the original

the non-mouse-passaged strain PMSS1. In contrast to that, HopQI producing strain B8 is colonizing in Mongolian gerbils with an acute pathology. Therefore, CEACAMs might contribute to the observed differences regarding the pathology and colonization in animals. At least the primordial members of the CEACAM receptor family are present in relatively simple species as well as in highly developed ones, such as mammals and, for this reason, most likely are also present in Mongolian gerbils. They execute main cellular functions like cell-cell-adhesion and are involved in signaling of the innate and adaptive immunity. Although *H. pylori* did not adhere to several orthologues of CEACAM1, it is not utterly impossible that *H. pylori* is able to cross species barriers by binding to CEACAMs found in other species. The N-terminal binding domain of CEACAMs is exposed to low selective pressure and, therefore, highly variable in aa sequence between the species, but a proof for the impotence in binding is still missing.

Taken together, the results obtained so far from this study concerning the identification of HopQ as a novel bacterial adhesin and its interaction with CEACAM molecules have brought interesting new insights into *H. pylori* pathogenesis and might be the basis for a more precise understanding of the interplay between this pathogenic bacterium and its human host.

# Material

## 1 Cell lines

Cell line	Properties	Reference
AGS	human adenocarcinoma cell line, ATCC CRL-1739	Barranco <i>et al.</i> , 1983
HEK293	human embryonic kidney cell line, ATCC CRL-1573	Graham <i>et al.</i> , 1977
HEK293/CEACAM1	human CEACAM1 transfected HEK293	Nägele, Dissertation 2010
HEK293/CEACAM5	human CEACAM5 transfected HEK293	Nägele, Dissertation 2010
HeLa	human cervix carcinoma cell line, ATCC CCL-2	Scherer <i>et al.</i> , 1953
Kato III	human gastric carcinoma cell line, ATCC HTB-103	Sekiguchi <i>et al.</i> , 1978
MKN28	human adenocarcinoma cell line	Motoyama <i>et al.</i> , 1986
MKN45	human adenocarcinoma cell line	Motoyama <i>et al.</i> , 1986

**Table 1: List of cell lines.**

## 2 Bacterial strains

### 2.1 *E. COLI*

Strain	Description	Reference
DH5 $\alpha$	F- $\Phi$ 80d lacZ $\Delta$ M15 $\Delta$ (lacZYA-argF) U169 deoR recA1 endA1 hsdR17 (rK-, mK+) phoA supE44 $\lambda$ - thi-l gyr A96 relA1 (Invitrogen, Karlsruhe)	Hanahan, 1983
Top10	F-mcrA $\Delta$ (mrr-hsdRMS-mcrBC) $\Phi$ 80lacZ $\Delta$ M15 $\Delta$ lacO74 recA1 ara $\Delta$ 139 $\Delta$ (ara-leu)7697 galU galK rpsL (StrR) endA1 nupG (Invitrogen, Karlsruhe)	Grant <i>et al.</i> , 1990

**Table 2: *E. coli* that were used in this work.**

## 2.2 *H. PYLORI*

Strain	Properties	Reference
P12 strep	clinical isolate 888-0, streptomycin resistant	Schmitt and Haas, 1994
26695	wildtype strain	Tomb <i>et al.</i> , 1997
B8	gerbil adapted	Farnbacher <i>et al.</i> , 2010
J99 strep	streptomycin resistant wildtype strain	R. A. Alm <i>et al.</i> , 1999
Tx30a	ATCC 51932	Tummuru MK <i>et al.</i> , 1993
X47	mouse adapted	H. Kleanthous <i>et al.</i> , 2001
SS1	murine passaged isolate of 10700	A. Lee <i>et al.</i> , 1997
PMSS1	clinical isolate 10700	A. Lee <i>et al.</i> , 1997
P8 strep	clinical isolate 196A, streptomycin resistant	Hofreuter <i>et al.</i> , 1998
ATCC 43526	wildtype strain	Melchers K, <i>et al.</i> , 1996
VK H76	P12 $\Delta$ <i>cagI</i>	Fischer <i>et al.</i> , 2001
VK H77	P12 $\Delta$ <i>babA</i>	this work
VK H78	P12 $\Delta$ <i>babB1+2</i>	this work
VK H86	P12 $\Delta$ <i>sabB1+2</i>	this work
VK H26	P12 $\Delta$ <i>hopZ</i>	this work
VK H80	P12 $\Delta$ <i>hopQ</i>	this work
SOP306	26695 $\Delta$ <i>hopQ</i>	Odenbreit <i>et al.</i> , 2002
VK H114	B8 $\Delta$ <i>hopQ</i>	this work
VK H109	P12 $\Delta$ <i>hopQ::hopQI</i>	this work
VK H90	P12 $\Delta$ <i>hopQ::hopQII</i>	this work
<i>H. mustelae</i>	<i>Helicobacter</i> species	Solnick JV <i>et al.</i> , 2001
<i>H. nemestrinae</i>	<i>Helicobacter</i> species	Solnick JV <i>et al.</i> , 2001

**Table 3:** *H. pylori* strains that were used in this work.

## 2.3 OTHER BACTERIA

Strain	Properties	Reference
<i>Moraxella catarrhalis</i> ATCC 43617	CEACAM-binding negative	Voges <i>et al.</i> , 2010
<i>Moraxella catarrhalis</i> ATCC 25238	CEACAM-binding positive	Voges <i>et al.</i> , 2010
<i>Neisseria gonorrhoeae</i> N302	CEACAM-binding negative	Kuespert <i>et al.</i> , 2007
<i>Neisseria gonorrhoeae</i> N309	MS11 strain constitutively producing Opa <sub>52</sub> , CEACAM-binding positive	Kuespert <i>et al.</i> , 2007

**Table 4:** *M. catarrhalis* and *N. gonorrhoeae*.

### 3 Plasmids

Name	Properties	Reference
pBluescript II SK (+)		Agilent Technologies
pUC19		Life Technologies
pSR11	pBluescript II SK (+) + <i>rpsL-erm</i> -cassette	lab collection, Stefanie Rohrer
pVK1+ <i>rpsLerm</i>	pSR11 + flanking regions of <i>hopZ</i> (P12)	this work
pVK1- <i>rpsLerm</i>	pBluescript II SK (+)+ flanking regions of <i>hopZ</i> (P12)	this work
pVK5+ <i>rpsLerm</i>	pSR11 + flanking regions of <i>sabB1</i> (P12)	this work
pVK5- <i>rpsLerm</i>	pBluescript II SK (+)+ flanking regions of <i>sabB1</i> (P12)	this work
pVK6+ <i>rpsLerm</i>	pSR11 + flanking regions of <i>sabB2</i> (P12)	this work
pVK6- <i>rpsLerm</i>	pBluescript II SK (+)+ flanking regions of <i>sabB2</i> (P12)	this work
pVK13+ <i>rpsLcat</i>	pUC19 + <i>dif-rpsL-cat-dif</i> -cassette + flanking regions of <i>hopZ</i> (P12)	this work
pIB6	<i>E. coli</i> - <i>H. pylori</i> <i>pHel3</i> shuttle plasmid	lab collection, Iris Barwig
pCE39	pIB6 + <i>hopQI</i> (P12)	lab collection, Claudia Ertl
pCE42	pIB6 + <i>hopQII</i> (Tx30a)	lab collection, Claudia Ertl

**Table 5: List of plasmids generated or used in this work**

## 4 Oligonucleotides and siRNA

All oligonucleotides were ordered from the company biomers.net. The table lists all primers used in this work. Restriction sites are underlined and tags are in bold.

Primer	Sequence 5' → 3'	Comment
VK1	GATCCTCGAGCCAGGCTTTTAGATAATTTG	fwd <i>hopZ</i> (P12) upstream, XhoI
VK2	GATCATCGATGTTAAACCCTTTGTGAAAC	rev <i>hopZ</i> (P12) upstream, ClaI
VK3	GATCGCGGCCGCTTTGAAATTAATTGAGTG	fwd <i>hopZ</i> (P12) downstream, NotI
VK4	GATCCC GCGGGGCTTGAGTCTAGAAAAC	rev <i>hopZ</i> (P12) downstream, SacII
VK9	GATCCTCGAGTTTAGACAACAAGGAATTGG	fwd <i>sabB1</i> (P12) upstream, XhoI
VK10	GATCATCGATTTGTAAACATGAATAGAAATG	rev <i>sabB1</i> (P12) upstream, ClaI
VK11	GATCGCGGCCGCGCTTCATCAAATCTATTTTG	fwd <i>sabB1</i> (P12) downstream, NotI
VK12	GATCCC GCGGAAGTGGTCCATAATGAG	rev <i>sabB1</i> (P12) downstream, SacII
VK13	GATCCTCGAGTATAGTATAGTAAAACATCG	fwd <i>sabB2</i> (P12) upstream, XhoI
VK14	GATCATCGATGAGTGATTCAAGCTCTC	rev <i>sabB2</i> (P12) upstream, ClaI
VK15	GATCGCGGCCGCTAACTATTTATCTTTAGAGTG	fwd <i>sabB2</i> (P12) downstream, NotI
VK16	GATCCC GCGCAAGAAATGTTGCAAAGTG	rev <i>sabB2</i> (P12) downstream, SacII
VK23	GATCCTCGAGTTAACGGGCTTAAGAATTGG	fwd <i>babB2</i> (P12) upstream, XhoI
VK24	GATCATCGATCTTAAATCCCTTTGTGGAAC	rev <i>babB2</i> (P12) upstream, ClaI
VK25	GATCGCGGCCGCTTTTGGATTAAATTAGG	fwd <i>babB2</i> (P12) downstream, NotI
VK26	GATCCC GCGTTGAACTAAGGAGAATGC	rev <i>babB2</i> (P12) downstream, SacII
VK32	GTAACATGGAAGACATGGATG	fwd <i>sabB1</i> (P12) flanking
VK33	TTGTCAGCTACGCGTTCTTG	fwd <i>sabB2</i> (P12) flanking
VK35	GATCTAACAAATCTTATTAGG	fwd <i>hopZ</i> (P12) flanking
VK36	AGTGGGCTATGAAATCATC	rev <i>hopZ</i> (P12) flanking
VK47	GGGTGCTTTTGCAACTCGC	rev <i>sabB1</i> (P12) flanking
VK48	TTTTGCGAGCGTTTCGGCG	rev <i>sabB2</i> (P12) flanking
VK57	GATAGTCGACAGGCGAGAATGGAAGTGATTAAG	fwd <i>hopQI</i> (P12) upstream, Sall
VK58	TGATTGAACGCGGATCCTTTAAGGTATAATGTT TCCGCCAC	rev <i>hopQI</i> (P12) upstream, BamHI
VK59	TATACCTTAAAGGATCCGCGTTCAATCAAGCTT GTAAGAG	fwd <i>hopQI</i> (P12) downstream, BamHI
VK60	GATCGAATTCTCCTTTAAGGGTGAGCTACATTG	rev <i>hopQI</i> (P12) downstream, EcoRI
VK63	TAATACGACTCACTATAGGG	<i>T7</i> forward primer
VK64	TAGAAGGCACAGTCGAGG	<i>BGH poly A</i> reverse primer
VK69	GATCGAATTCAGACTACAATCAATTGCGGTG	Sequencing <i>hopQI</i> (PMSS1), EcoRI
VK70	GATCGGTACCCATGTATTGGGTTTTTTTTGC	fwd <i>oipA</i> (P12) upstream, KpnI
VK71	GATCCCCGGGCATGCATCAAACAATAAAAC	rev <i>oipA</i> (P12) upstream, SmaI
VK72	GATCGCGGCCGCGCTTTTGATTGCTTGTC	fwd <i>oipA</i> (P12) downstream, NotI
VK73	GATCCC GCGGAAGCGTTTGAGGTTAAATCG	rev <i>oipA</i> (P12) downstream, SacII
VK74	CTTCTTCTATAGGCACATTCG	fwd <i>oipA</i> (P12) flanking

VK75	ATTTAGGCTCAAACACGGCG	rev <i>oipA</i> (P12) flanking
VK82	TTAAAAGCTAGCCAATCAAGTG	Sequencing <i>hopQ</i> (PMSS1, SS1)
CE61	CAATCAATTGCGGTGGGAG	Sequencing <i>hopQI</i> (P12) allele
CE62	ACTTTTTACAACCAGCCAG	Sequencing <i>hopQII</i> (Tx30a) allele
CE67	GATCGTCGACATGTACCCATACGATGTTCCAGA TTACGCTATGAAAAAACGAAAAAAC	fwd <i>hopQI</i> (P12) and <i>II</i> (Tx30a), SalI, HA-tag
CE68	GATCGTCGACATGAAAAAACGAAAAAAC	fwd <i>hopQI</i> (P12) and <i>II</i> (Tx30a), SalI
CE69	GATCAGATCTTTTAATACGCGAACACATAA	rev, <i>hopQI</i> (P12), BglII
CE70	GATCAGATCTTTTAATAGGCAAACACATAA	rev, <i>hopQII</i> (Tx30a), BglII
CE82	TGGTGATAAAGGTCGTTAAACCCGC	fwd <i>hopQI</i> (P12) flanking
CE83	CGGCGATGGAATAAACTCTAAGGC	rev <i>hopQI</i> (P12) flanking
LH3	GATCCTCGAGGTAGTTGGTTTAAGCGGTTG	fwd <i>babA</i> upstream, XhoI
LH4	GATCATCGATCTTAAATCCCTTTGTGGAAC	rev <i>babA</i> upstream, ClaI
LH5	GATCGCGGCCGCATTAGTATTAGGGATTTTAC	fwd <i>babA</i> downstream, NotI
LH6	GATCCCGCGGTATCGTTACAAGCGCATTTG	rev <i>babA</i> downstream, SacII
LH7	GATCCTCGAGATGCCGGCATTAGTAAAAAG	fwd <i>babB1</i> upstream, XhoI
LH8	GATCATCGATCTTAAATCCCTTTGTGGAAC	rev <i>babB1</i> upstream, ClaI
LH9	GATCCCGCGGCCGCGAGAGAGTAAAAGGGTTTTTC	fwd <i>babB1</i> downstream, NotI
LH10	GATCCCGCGGCAAACGCTATGCAAGATGG	rev <i>babB1</i> downstream, SacII
SR21	CCGGATCCGTGATGCGGTGCG	RAPD PCR, D9355
SR42	CAGCATCGATATGCTTTATAACTATGGATT	fwd <i>rpsL<sup>S</sup>-erm<sup>R</sup></i> -cassette, ClaI
SR43	GCGGTACCTTACTTATTAATAATTTATAG	rev <i>rpsL<sup>S</sup>-erm<sup>R</sup></i> -cassette, KpnI
WS614	CAGGGTTATACGCCACGGCGATG	rev <i>hopQI</i> (P12) flanking

**Table 6: List of oligonucleotides.**

The siRNA was obtained from Qiagen and used in the final concentration of 10 $\mu$ M:

siRNA	Catalog number	Target sequence 5' → 3'	Gene
Hs_CEACAM1_11	SI05459846	CTGCACAGTACTCCTGGCTTA	CEACAM1, 5 and 6
Allstars Hs Cell Death Control	SI04381048	not specified	plk (polo-like kinase)
Allstars Negative Control	SI03650318	not specified	--

**Table 7: List of siRNA.**

## 5 Commercially available kits

Name	Description	Supplier
QIAprep Spin Miniprep Kit	Isolation of plasmid DNA	Qiagen
QIAamp Tissue Kit	Isolation of genomic DNA	Qiagen
illustra GFX PCR DNA and Gel Band Purification Kit	Purification of DNA from agarose gels or PCR reactions	GE Healthcare
Wizard® Plus SV Minipreps DNA Purification System	Isolation of plasmid DNA from <i>H. pylori</i>	Promega

**Table 8: List of commercially available kits.**

## 6 Antibodies

Name	Description	Supplier
$\alpha$ -CagA (AK268)	polyclonal antibody against N-terminal part of CagA ( <i>H. pylori</i> 185-44), rabbit	W. Fischer
$\alpha$ -GFP (JL-8)	living colors A.v. monoclonal antibody, clone JL-8, mouse	Clontech
$\alpha$ -HopQI	polyclonal peptide antibody (SKGEKLEAHVTTSKYQC) of <i>H. pylori</i> P12	Genosphere Biotech, this work
$\alpha$ -OipA (AK282)	polyclonal antibody, rabbit	S. Odenbreit
$\alpha$ -P-Tyr (4G10)	monoclonal antibody against tyrosine phosphorylation, clone 4G10, mouse	Upstate Millipore
$\alpha$ -RecA (AK263)	polyclonal antibody against RecA of <i>H. pylori</i> P1, rabbit	Schmitt <i>et al.</i> , 1995
$\alpha$ -tubulin	monoclonal, mouse	Abcam
pan-CEACAM	monoclonal, clone D14HD11, recognizes CEACAM1, 3, 4, 5 and 6, mouse	Aldevron Freiburg
$\alpha$ -CEACAM1	monoclonal, clone GM-8G5, mouse	Aldevron Freiburg
$\alpha$ -CEACAM5	monoclonal, clone 26/3/13, mouse	Aldevron Freiburg
$\alpha$ -CEACAM6	monoclonal, clone 9A6, mouse	Aldevron Freiburg
$\alpha$ -rabbit IgG-HRP	horse radish peroxidase polyclonal antibody against rabbit IgG, goat	Santa Cruz
Protein A-AP	alkaline phosphatase polyclonal antibody IgG	Sigma-Aldrich
Alexa Fluor®488 goat $\alpha$ -mouse IgG	monoclonal antibody, $\alpha$ -mouse, labelled with Alexa Fluor® 488 dye	Life Technologies
Alexa Fluor®488 goat $\alpha$ -rabbit IgG	monoclonal antibody, $\alpha$ -rabbit, labelled with Alexa Fluor® 488 dye	Life Technologies
Interleukin-8 - biotin- IL-8 (human)	ELISA, mouse IgG2b, coating	BD PharMingen
Interleukin-8 - IL-8 (human)	ELISA, mouse IgG2b, detection antibody	BD PharMingen

**Table 9: List of antibodies.**

## 7 Bacterial culture medium

Name	Composition	Supplier
LB liquid media	20 g/l Lennox-L-medium	Gibco/Invitrogen
LB agar plates	32 g/l Lennox-L-Agar	Gibco/Invitrogen
BB liquid media	28 g/l Brucella Broth	Falcon BD
GC agar plates	36 g/l GC-Agar-Base	Oxoid
Serum plates	GC agar base, supplemented with 8% horse serum, 5 µg/ml trimethoprim, 1% vitaminmix and 1 µg/ml nystatin after sterilization by autoclave	

**Table 10: List of culture medium for bacterial growth.**

## 8 Buffers and solutions

Name	Composition	Application
GEBS	20% (v/v) glycerol, 50 mM EDTA, 0.05% (w/v) Bromophenol blue, 0.5% (w/v) N-laurylsarcosyl	Agarose gels
TAE-buffer	40 mM Tris, 20 mM acetic acid, 1 mM EDTA	Agarose gels
SOC medium	2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl <sub>2</sub> , 10 mM MgSO <sub>4</sub> , and 20 mM glucose	<i>E. coli</i> transformation, cloning
Electroporation buffer	272 mM sucrose, 15% (v/v) glycerol; 2.43 mM K <sub>2</sub> HPO <sub>4</sub> ; 0.57 mM KH <sub>2</sub> PO <sub>4</sub>	Electroporation
TfbI buffer	30 mM potassium acetate, 100 mM RbCl, 10 mM CaCl <sub>2</sub> , 50 mM MnCl <sub>2</sub> , 15% (v/v) glycerol; pH 5.2 by titration with 0.2 M acetic acid; sterile filtrated and kept on ice	Generation of competent cells
TfbII buffer	10 mM MOPS, 75 mM CaCl <sub>2</sub> , 10 mM RbCl; 15% (v/v) glycerol; pH 6.5 by titration with KOH; sterile filtrated and kept on ice	Generation of competent cells
Coating buffer	100 mM Na <sub>2</sub> HPO <sub>4</sub> pH 9.6	IL-8 ELISA
Washing buffer	0.05% (v/v) Tween-20 in PBS	IL-8 ELISA
Blocking buffer	10% FCS in PBS	IL-8 ELISA
PBS*	PBS, 1 mM EDTA, 1 mM sodium vanadate, 1 mM PMSF, 1 µM leupeptin, 1 µM pepstatin	In vitro infection

2xSDS loading buffer	50 mM Tris/HCl, pH 6.8; 100 mM DTT; 2% (w/v) SDS; 10% (w/v) glycerol; 5% (v/v) $\beta$ -mercaptoethanol; 0.1% (w/v) Bromophenol blue	SDS PAGE
Coomassie solution	0.275% (w/v) Coomassie Brilliant Blue R 250 in 50% methanol, 10% (v/v) acetic acid	SDS PAGE
Coomassie destaining	10% (v/v) methanol, 10% (v/v) ethanol, 7.5% (v/v) acetic acid	SDS PAGE
Electrophoresis buffer	250 mM glycine, 0.1% (w/v) SDS, 25 mM Tris/HCl pH 8.3	SDS PAGE
2x single-gel system buffer	152 mM Tris/HCl, 0.2 M serine, 0.2 M glycerol, 0.2 M asparagine, pH 7.4; sterile filtrated	SDS PAGE
Alkaline phosphatase development solution	0.1 M Tris/HCl pH 9.6, 0.1 g/l NBT, 7 mM MgCl <sub>2</sub> , 50 mg/l BCIP	Western Blot
Stripping solution	25 mM glycine-HCl, pH = 2, 1% (w/v) SDS	Western Blot
Western transfer buffer	192 mM glycine, 25 mM Tris, 20% (v/v) methanol, 0.1% (w/v) SDS, pH = 8.3	Western Blot
Blocking solution	3% (w/v) BSA in TBS	Western Blot
Washing buffer TBST	0.075% (v/v) Tween-20 in TBS	Western Blot
Vitamin mix	100 g/l d-glucose, 10 g/l l-glutamine, 26 g/l l-cystein, 0.1 g/l cocarboxylase, 20 g/l FE(III) nitrate, 3 g/l thiamin, 13 mg/l p-aminobenzoic acid, 250 mg/l nicotinamide-adenine dinucleotide, 10 mg/l guanine, 0.15 g/l l-arginine, 0.5% uracil	agar plates ( <i>H. pylori</i> )
PBS+	0.9 mM CaCl <sub>2</sub> , 0.5 mM MgCl <sub>2</sub> in PBS	Bacterial pull-down Assay
FACS buffer	2% FCS in PBS	Bacterial pull-down assay, Cell staining

**Table 11: List of buffers and solutions.**

---

## 9 Chemicals and enzymes

All chemicals were ordered in the purity level pro analysis mainly from the companies Merck, Roth and Sigma-Aldrich. Restriction enzymes were obtained from Roche and Fisher Scientific. The supplying company of any other used enzymes and chemicals is mentioned directly in the text at the appropriate passage.

## 10 Consumables

Adhesive Plate Seals (ABgene), cover slips (Menzel), cryogenic vials (Nalgene), Omnifix® syringes 5 ml and 50 ml (Braun), electroporation cuvettes (Peqlab), Eppendorf reaction cups 1.5 ml and 2ml (Eppendorf), x-ray film cassettes (Kodak), Falcon vials 15ml and 50ml (Falcon), Whatman paper (Whatman), Gas permeable adhesive seals (ABgene), Spectrophotometry cuvettes (Brand), pipettes and tips (Gilson), Petri dishes (Greiner), PVDF membrane (Bio-Rad), Nitrocellulose membrane filter (Millipore), slide (Langenbrinck), x-ray films Super RX (Fuji Film), Quick-Seal™ vials (Beckmann Coulter Inc.), cell culture flasks (Falcon), 0,2 µm syringe filter (Josef Peske GmbH), Nunc cell culture flasks (Thermo scientific) cell scraper (Falcon), 6, 12, 24, 48 and 96-well plates (Thermo scientific), Nunc™ 96-Well Polystyrene Round Bottom Microwell Plates (Thermo scientific)

## 11 Equipment and devices

Agarose gel chambers Wide Mini-Sub® cell (Bio-Rad), anaerobic jars (Fritz Gössner GmbH) Anaerobic bacteria microincubator MI22C (Scholzen), incubator FED (Binder), Flow Cytometer Canto II (BD), Gene Pulser Xcell™ Electroporation System (Bio-Rad), -80°C freezer (Heraeus), Gel Doc 2000 System Quantity One 4.4.0 (Bio-Rad), Electrophoresis system Mini-Protean III™ (Bio-Rad), heat block (Techne), fridge (Liebherr), magnetic stirrer with heating system MR 3001 (Heidolph), microwave (AEG), pH meter ProfiLine pH 197i (WTW), photometer Libra S12 (Biochrom), rotating mixers (Assistent), shaker (Eppendorf), power supply Power Pac 300 (Bio-Rad), working bench (BDK Luft-und Reinraumtechnik GmbH), thermocycler Microcycler Personal (Perkin Elmer), ultrasound device Sonifier II 450 (G. Heinemann), table centrifuges

(Megafuge 1.0R, Biofuge 15R, Biofuge 15) (Heraeus), Vortex Gene 2 (Scientific Industries), scales (Biotech Fischer), water bath 1012 (GFL), semi-dry western blotting system (Biotech Fischer), centrifuge RC5C Plus (Sorvall), cyclor peqSTAR 96 universal gradient (Isogen Life Science)

# Methods

---

## 1 Working with eukaryotic cells

### 1.1 MAINTENANCE OF CELL LINES

Eucaryotic cell lines were grown at 37°C, 5% CO<sub>2</sub> and subcultured every three to four days. AGS and HeLa cells were grown in RPMI/10% FCS, MKN45 cells in RPMI/20% FCS and Kato III cells in DMEM/20% FCS. For MKN28 RPMI medium was supplemented with 10% FCS, 10mM Hepes, 1 mM sodium pyruvate and 0.1 mM nonessential aa. HEK293 cells were cultured in DMEM/15% FCS and supplemented with 1 mg/ml G418, when they were transfected with CEACAM constructs.

### 1.2 CRYOCONSERVATION

1·10<sup>6</sup> cells were detached from the cell culture flask by incubating with 2 mM EDTA/PBS for 10 min and resuspended in medium/10% FCS. After centrifugation at 900 g for 5 min, the supernatant was discarded and the cells were dissolved in 1 ml FCS/10% DMSO, transferred into a cryogenic vial and stored in liquid nitrogen.

### 1.3 CELL STAINING

All centrifugation steps were carried out at 4°C and 900 g for 5 min.

Adherent cells were first detached from the cell culture flasks by adding 2 mM EDTA/PBS for 15 min at 37°C, resuspended in PBS/10% FCS and seeded in 96-well plates. Suspension cell lines were centrifuged at 900 g for 5 min, resuspended in PBS and then seeded in 96-well plates.

Cells were washed three times with ice cold PBS and then incubated with primary antibody for 1 hour at 4°C. After three further washing steps with ice cold PBS, cells were incubated with the secondary antibody Alexa Fluor®488 goat α-mouse IgG for 1 hour at 4°C. The cells were washed again and resuspended in PBS/2% FCS and analyzed by flow cytometry.

The obtained MFI values were normalized against the negative control that was a sample of cells only incubated with the secondary antibody and set to the value of 1. The fold production level

was determined by comparing the MFI values of each sample to the MFI value of the negative control.

#### 1.4 KNOCKDOWN BY RNA INTERFERENCE

AGS cells were transfected in 12-well plates with a final siRNA concentration of 10  $\mu$ M. Lipofectamine RNAiMAX (Invitrogen) was used as transfection reagent according to the manufacturer's recommendations. The siRNA effect on synthesis of CEACAM1, 5 and 6 was evaluated after 48 hours against AGS control cells by flow cytometry analysis.

## 2 Working with bacteria

### 2.1 CULTIVATION OF BACTERIA

#### 2.1.1 *E. COLI*

*E. coli* were cultured on LB<sub>0</sub>-agar plates or in LB liquid media under aerobic conditions at 37°C. For the selection of positively transformed bacteria, LB plates were supplemented with ampicillin (100 µg/ml), chloramphenicol (30 µg/ml), kanamycin (50 µg/ml), erythromycin (250 µg/ml) or streptomycin (250 µg/ml).

For cryoconservation, bacteria were resuspended in 1 ml of LB medium/20% glycerol in cryogenic vials and stored at -80°C.

#### 2.1.2 *H. PYLORI*

*H. pylori* strains were grown at 37°C under microaerophilic conditions on GC agar plates supplemented with horse serum (8%), trimethoprim (5 µg/ml), vitaminmix (1%) and nystatin (1 µg/ml) (serum plates). Chloramphenicol (6 µg/ml), kanamycin (8 µg/ml), erythromycin (10 µg/ml) or streptomycin (250 µg/ml) was added to select for transformants or screen colonies for resistance to either drug. Plates were incubated for 24 h under microaerobic conditions (85% N<sub>2</sub>, 10% CO<sub>2</sub>, 5% O<sub>2</sub>) at 37°C.

Bacteria were passaged 2 to 3 times before the first experiments were performed.

For cryoconservation, bacteria were resuspended in 1 ml of Brucella Broth (BB) medium supplemented with 20% glycerol and 10% FCS in cryogenic vials and stored at -80°C.

#### 2.1.3 OTHER BACTERIA

*M. catarrhalis* strains ATCC 43617 and ATCC 25238 were incubated at 37°C for 24 h using Columbia blood agar plates.

*N. gonorrhoeae* strains N302 and N309 were cultured at 37°C and 10% CO<sub>2</sub> on GC plates supplemented with horse serum (8%), and vitamin mix (1%).

Bacteria were passaged 2 to 3 times before the first experiments were performed.

For cryoconservation, bacteria were resuspended in 1 ml of LB medium/20% glycerol in cryogenic vials and stored at -80°C.

## 2.2 GENERATION OF CHEMICALLY COMPETENT *E. COLI* (TOP 10)

A frozen stock of chemically competent Top 10 cells was used for over night inoculation of a 10 ml LB-liquid culture at 26°C and 180 rpm. At an OD<sub>550</sub> of 0.6, the cells were transferred into a 200 ml liquid culture and incubated at 37°C, until they reached the optimal OD<sub>550</sub> of 0.56. In the meantime, the buffers TfbI and TfbII were freshly prepared and kept cold. The lids of 1.5 ml tubes were closed in advance and the tubes prefrozen at -20°C.

When the OD<sub>550</sub> was reached, the culture was chilled down for 5 min on ice and centrifuged at 3000 rpm for 15 min at 4°C. The supernatant was discarded and the cells were carefully resuspended in 80 ml of TfbI buffer. Cells were again incubated for 5 min on ice, centrifuged at 3000 rpm for 15 min at 4°C and resuspended in 8 ml of TfbII buffer. Immediately after, aliquots of 50 µl were filled into the cold 1.5 ml tubes and shock-frozen with liquid nitrogen.

## 2.3 GENERATION OF *H. PYLORI* KNOCKOUT MUTANTS

For the generation of markerfree outer membrane protein mutants in *H. pylori*, the method of contra-selectable streptomycin susceptibility was used<sup>135</sup>. For this, a streptomycin resistant strain (*rpsL<sup>R</sup>*) as starting point was essential. If the wildtype strain was not streptomycin resistant from the beginning, the strain was made by transformation with the plasmid pEG21.

About 1000 bp of the up and downstream regions of the gene to be deleted were cloned in a pBluescript II SK (+) vector with a resistance cassette, referred to as *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette (Deletion plasmid A) and a pBluescript II SK (+) vector without the resistance cassette, respectively (Deletion plasmid B). Primers used for the amplification of the up and downstream regions of the distinct outer membrane proteins are listed in Table 6.

In the first step, the generated deletion plasmid A with the *rpsL<sup>S</sup>-erm<sup>R</sup>*-cassette was used for transformation of the streptomycin resistant wildtype strain. By homologous recombination of the adjacent gene regions, the resistance marker replaced the gene to be deleted and the distinct omp mutant was now erythromycin resistant and streptomycin sensitive.

In the second step, the resistance cassette was completely removed by transformation with the deletion plasmid B resulting in a markerfree mutant strain that was resistant for streptomycin again.

## 2.4 TRANSFORMATION OF *H. PYLORI*

Bacteria of an over night culture were solved in 1 ml PBS and the OD<sub>550</sub> was measured.  $6 \cdot 10^7$  bacteria (OD<sub>550</sub> = 0.2) were centrifuged for 5 min at 4000 rpm and the supernatant was discarded. The bacteria were resuspended in 1 ml BB/10% FCS and incubated for 1 or 2 hours at 37°C, 10% CO<sub>2</sub>. Next, 500 to 1000 ng plasmid DNA (isolated from *E. coli*) or half the amount for DNA isolated from *H. pylori* was added to the transformation. The DNA-bacteria mixture was then incubated for further 4 to 6 hours at 37°C, until it was centrifuged for 5 min at 4000 rpm, resuspended in 100 µl BB medium and plated onto a selective agar plate. After incubation for 3 to 4 days under micro aerobic conditions at 37°C, the transformation was screened for positive colonies.

## 2.5 ELECTROPORATION OF *H. PYLORI*

Bacteria of an over night culture were dissolved in 1 ml PBS and the OD<sub>550</sub> was measured.  $3 \cdot 10^8$  bacteria (OD<sub>550</sub> = 1) were centrifuged for 5 min at 4000 rpm and the supernatant was discarded. All steps from here on were carried out at 4°C or on ice and all reagents were kept cool. The bacteria were resuspended in 500 µl ice cold electroporation buffer and centrifuged for 5 min at 4000 rpm. This step was repeated twice. The bacteria were then solved in 55 µl of electroporation buffer and 200 to 500 ng plasmid DNA (isolated from *E. coli*) was added to the mix. If the plasmid was isolated from *H. pylori* or genomic DNA was used, less amounts of DNA were sufficient for an efficient electroporation. The DNA-bacteria mixture was then transferred to a precooled electroporation cuvette (PeqLab) and the electroporation was performed at 2.5 kV, 200 Ω, 25 µF. Immediately after the pulse, 1 ml BB/10%FCS was added to the cuvette. After passaging the mixture carefully back and forth between the cuvette and a 1.5 ml reaction cup, the mixture was transferred to a 24-well plate and incubated over night at 37°C, 10% CO<sub>2</sub>. The next day, the mixture was transferred into a 1.5 ml reaction cup and centrifuged for 5 min at 4000 rpm. The pellet was solved in 100 µl BB medium and plated onto a selective agar plate. After incubation for 3 to 4 days under micro aerobic conditions at 37°C, the electroporation was screened for positive colonies.

## 2.6 BACTERIAL PULL-DOWN ASSAY

The bacterial pull-down was performed as described before (4). Bacteria of an overnight culture were resuspended in 1 ml PBS and the OD<sub>550</sub> was measured.  $4 \cdot 10^6$  bacteria, as calculated by the OD<sub>550</sub> measurement, were incubated with cell culture supernatants containing CEACAM-N-GFP. After an incubation time of 30 min at room temperature with head-over-head rotation, the bacteria were washed twice with PBS containing 0.9 mM CaCl<sub>2</sub> and 0.5 mM MgCl<sub>2</sub> and resuspended in SDS loading buffer. After heat denaturation the samples were analyzed by western blotting. Alternatively, bacteria were resuspended in PBS/2% FCS and analyzed by flow cytometry (FACS Canto II, BD Biosciences).

The obtained MFI values were normalized against the negative control that was a sample of bacteria alone set to the value of 1. The fold binding was determined by comparing the MFI values of each sample to the MFI value of the mock control.

## 2.7 *IN VITRO* INFECTION EXPERIMENTS WITH *H. PYLORI*

The day before infection, eukaryotic cell lines were seeded in 6-well plates so they would reach a confluency of 80 to 90% the next day and serum-starved over night using cell culture medium without FCS. Two hours prior to infection, the medium was changed to 1 ml PBS/10% FCS per 6-well.

Bacteria of an overnight culture were resuspended in 1 ml PBS and the OD<sub>550</sub> was measured. In all experiments, a MOI (multiplicity of infection) of 60 was used and the infection was stopped after 4 hours on ice. The supernatant was stored at -20°C and analyzed later for production of IL-8. Cells were harvested in ice cold PBS containing 1 mM sodium vanadate, 1 μM leupeptin, 1 μM pepstatin and 1 mM PMSF (PBS\*) and resuspended in 2xSDS loading buffer. Equal amounts of samples were loaded on a gel and analyzed by western blotting.

Tubulin was used as control that equal amounts of cells were analyzed during the infection with *H. pylori*. AK268 was directed against the N-terminal part of CagA and stained bacterial CagA. Antibody 4G10 showed the level of CagA phosphorylation e.g. the amount of CagA translocated into the host cells and tyrosine phosphorylated by cellular kinases.

## 3 DNA methods

### 3.1 ISOLATION OF DNA

#### 3.1.1 PLASMIDS

For the screening of positive clones, plasmid DNA from an overnight culture of *E. coli* was isolated using the QIAprep Spin Miniprep Kit (Qiagen) and finally eluted in 25 µl of TE-buffer.

Plasmid DNA from *H. pylori* was gained by using the Wizard™ Plus SV Miniprep Kit (Promega) according to the manufacturer's recommendations and eluted in 20 µl of DNase free water. Plasmid DNA was stored at -20°C.

#### 3.1.2 GENOMIC DNA

Genomic DNA of *H. pylori* was isolated from bacterial overnight cultures. About ¼ of an agar plate was resuspended in PBS and centrifuged for 5 min at 6000 g. After discarding the supernatant, the DNA was isolated by following the instructions of the QIAamp Tissue Kit (Qiagen), eluted in 100 µl of AE-buffer and stored at -20°C.

### 3.2 DNA CONCENTRATION AND QUALITY

The concentration of plasmid and genomic DNA was evaluated by using a Nanodrop (Peqlab). 1 µl of DNA was pipetted on the petals of the machine. The software measured the OD at 230, 260 and 280nm and helped to determine protein and salt contaminations.

A DNA sample of 200 µg was loaded on a 1% (w/v) agarose gel and run for 30 min at 110 V in TAE-buffer to test for DNA degradation. GeneRuler™ 1 kb DNA Ladder (Fermentas) was loaded on the gel to estimate the fragment size of the DNA.

The gel was incubated for 15 min in a bath with ethidium bromide (1mg/l) and analyzed under UV light at 260 nm (Molecular Imager Gel Doc XR System, Bio-Rad).

### 3.3 DNA EXTRACTION FROM AGAROSE GELS AND ENZYMATIC REACTIONS

DNA of the correct size was cut out from agarose gels under UV light and incubated at 60°C. The dissolved DNA of agarose gels was then extracted the same way as DNA from enzymatic reactions with the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) following the manufacturer's instructions. The DNA was eluted in 12 to 15 µl of Elution buffer type 4 (10 mM Tris-HCl, pH 8.0) and stored at -20°C.

### 3.4 AMPLIFICATION OF DNA FRAGMENTS

The method of polymerase chain reaction (PCR) was used to amplify DNA fragments from template DNA. For this, specific oligonucleotides were designed that bound to short complementary sequences on the plus and minus strand of the template DNA, respectively. By repeating the three steps of PCR, - DNA double strand denaturation, oligonucleotide annealing to the two single DNA strands and elongation of oligonucleotides -, in each cycle two new double strands of DNA were produced.

#### 3.4.1 PCR UNDER STANDARD CONDITIONS

All PCR reactions were performed in the cycler peqSTAR 96 universal gradient (Isogen Life Science). Either plasmid DNA (100-250 ng) or genomic DNA (1-10 ng) was used as template DNA for the reaction mixture.

The Taq Polymerase from PAN Biotech GmbH was used for analytical purposes only. The polymerase had a 5'-3' exonuclease activity. The protocol for a standard 25 µl reaction mix is shown in Table 12.

Volume	Reagents
2.5 µl	10x PCR buffer without MgCl <sub>2</sub>
2.5 µl	10 mM dNTP mix
1.5 µl	50 mM MgCl <sub>2</sub>
1 µl	10 µM forward oligonucleotide
1 µl	10 µM reverse oligonucleotide
0.1 µl	Taq Polymerase PAN
ad 25 µl	H <sub>2</sub> O bidest/DNA template

**Table 12: Standard reaction conditions Taq polymerase (PAN Biotech GmbH).**

The Ex Taq DNA Polymerase from TaKaRa was applied to reactions, if the DNA fragments were later used in cloning experiments or for sequencing. The Ex Taq DNA Polymerase has a 3'-5' exonuclease activity that reduces the amount of wrongly inserted bases in the newly synthesized DNA fragment. The polymerase is suitable for DNA amplifications of sizes up to 20 kbp. The protocol for a 50 µl reaction mix is shown in Table 13.

Volume	Reagents
5 µl	10X Ex Taq Buffer
4 µl	2.5 mM dNTP mix
4 µl	25 mM MgCl <sub>2</sub>
1 µl	10 µM forward oligonucleotide
1 µl	10 µM reverse oligonucleotide
0.1 µl	Ex Taq DNA Polymerase
ad 50 µl	H <sub>2</sub> O bidest/DNA template

**Table 13: Standard reaction conditions Ex Taq DNA polymerase (TaKaRa).**

The standard PCR program used for either Ex Taq DNA polymerase (TaKaRa) or Taq polymerase (PAN Biotech GmbH) is shown in Table 14. The annealing temperature for the oligonucleotides was calculated by the software DNAMAN (Version 6). The elongation time for each PCR reaction was modified in regard to the length of the DNA fragment to be amplified. For every 1000 bp DNA fragment one minute elongation time was used. The result of the PCR was analyzed by agarose gel electrophoresis.

PCR steps	Temperature	Time	Repeats
First denaturation	94°C	4 min	
Denaturation	94°C	30 s	30x
Annealing	48 – 65°C	45 s	
Elongation	TaKaRa 68°C PAN 72°C	1 min/1000 bp	
Endelongation	TaKaRa 68°C PAN 72°C	10 min	
	16°C	∞	

**Table 14: Standard PCR program used for the DNA polymerases (TaKaRa/PAN).**

### 3.4.2 COLONY PCR (SAMBROOK ET AL, 1989)

For the screening of *E. coli* clones after transformation, a colony PCR was performed. That method gave rapid information about clones that had the correct insert or plasmid.

The Taq Polymerase (PAN Biotech GmbH) was used by default.

Clones on selective agar plates were picked with a tip and transferred to another selective plate. The same tip was further transferred into a labelled PCR reaction cup filled with 17 µl H<sub>2</sub>O bidest. After incubation of about 10 min, the tip was discarded and the water-clone-solution filled with the PAN PCR reagents. The standard PCR program for the PAN Taq polymerase was run. The result of the PCR was analyzed by agarose gel electrophoresis.

### 3.4.3 RAPD PCR

RAPD (randomly amplified polymorphic DNA) is a method for DNA fingerprinting. Genomic DNA can be assigned to strains of *H. pylori* regarding to a distinct band pattern that is produced by RAPD PCR and looks repeatedly the same for the same template DNA. The different DNA band patterns could be compared and identical patterns were used to ascertain the belonging to a specific strain. The conditions of the PCR varied from the standard conditions. Instead of two specific oligonucleotides, only one primer was added to the PCR mix (D9355). The amounts of genomic DNA, salt, dNTPs, oligonucleotide and polymerase were higher compared to the amount of a standard PCR mix (Table 15).

Volume	Reagents
2.5 µl	10X Ex Taq Buffer
3 µl	2.5 mM dNTP mix
3 µl	25 mM MgCl <sub>2</sub>
1 µl	30 µM D9355
0.25 µl	Ex Taq DNA Polymerase
	50 – 100 ng gDNA
ad 25 µl	H <sub>2</sub> O bidest

**Table 15: Standard reaction conditions RAPD PCR**

The PCR program was also different to the standard program. The annealing temperature was set to only 40°C in the first PCR cycles to allow the primer annealing to only partial complementary sequences. The DNA fragments generated by the oligonucleotide in the first cycles with low rigor were then amplified at a higher annealing temperature (**Table 16**). The result of the PCR was analyzed on a 2% agarose gel by agarose gel electrophoresis.

PCR steps	Temperature	Time	Repeats
Denaturation	94°C	5 min	4x
Annealing	40°C	5 min	
Elongation	72°C	5 min	
Denaturation	94°C	1 min	30x
Annealing	55°C	1 min	
Elongation	72°C	2 min	
Endelongation	72°C	10 min	
	16°C	∞	

**Table 16: RAPD PCR program**

## 3.5 CLONING

### 3.5.1 RESTRICTION OF DNA

All restriction enzymes were obtained from Roche or Thermo Fisher Scientific and used in analytical and preparative digestions as recommended by the manufacturer.

For analytical purposes, 100 to 300 ng plasmid DNA was digested by restriction endonucleases and incubated 1 to 2 hours, before the digestion with one or more enzymes was analyzed by agarose gel electrophoresis.

Preparative restrictions of plasmids were used for cloning experiments. About 700 to 800 ng plasmid DNA was digested and incubated for 3 to 4 hours or over night. The digestion was then purified using the illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) as described before (3.3).

If the buffers of two restriction endonucleases recommended by the manufacturer were not compatible, the restriction was done sequentially with a purification step in between the two reactions to make sure that both enzymes had a high activity and the DNA was digested properly. In this case, the loss of DNA during the purification step was compensated by using more plasmid DNA.

The amount of DNA after digestion was measurement by Nanodrop calculation (Peqlab).

To inactivate the restriction enzymes before ligation, the finished digestion mix was incubated for 20 min at 65°C.

### 3.5.2 LIGATION

The optimal amount of insert compared to vector is crucial for an efficient ligation. At best, the molar ratio between vector and insert DNA should be 1 to 5. Best results were achieved by using 25 ng of vector DNA and 5 times the amount of insert DNA. Since the molar ratio is important, the amount of insert DNA was calculated by the following formula:

$$m(\text{insert DNA}) = 5 \cdot 25\text{ng} (= \text{vector DNA}) \cdot \frac{\text{length insert DNA}}{\text{length vector DNA}}$$

Since the concentration of both insert and vector DNA were calculated after digestion by Nanodrop measurement, the needed volumes for the ligation reaction could be easily determined.

For sticky ends ligations, 1 µl 10x ligation buffer and 1 µl T4 DNA Ligase (Roche) was added to the insert and vector DNA-mix and incubated for 1 hour at 26°C. For difficult and blunt end ligations, more ligase (1 – 5 U) was added to the reaction mix and ligation was usually performed overnight in an ice bath. Also PEG (polyethylene glycol) was added to the blunt end mix according to the manufacturer's recommendations to increase the likelihood of the compatible ends being ligated.

### 3.5.3 TRANSFORMATION OF CHEMICALLY COMPETENT *E. COLI*

5 to 10 µl of the ligation mix was used for the transformation of chemically competent *E. coli* Top 10. 50 µl of bacteria were thawed on ice for 30 min, before the ligation mix was added to the bacteria. After 20 min incubation on ice, a heat shock was performed at 42°C for 45 s in a waterbath. The Top 10 were immediately after the heat shock put back on ice. After a few minutes incubation on ice, 500 to 1000 µl of SOC-medium were added to the transformation and the mix was shaken with 180 rpm for 1 to 2 hours at 37°C.

## 3.6 SEQUENCING

Sequencing was done by the company GATC Biotech AG (Kempten, Germany). Samples were prepared according to the company's instructions and the readout was evaluated by using the software CLC DNA Workbench version 6.0.2.

---

## 4 Protein methods

### 4.1 BACTERIAL LYSATES

Bacteria of an over night culture were solved in 1ml PBS and the OD<sub>550</sub> was measured. For an OD<sub>550</sub> of 0.1, the bacteria were centrifuged at 6000 g for 5 min and resuspended in 10 µl of PBS and 10µl of 2x SDS Loading Dye. The bacteria were then boiled for 10 min at 98°C to denature proteins by heat. For bacterial lysates 4 to 5 µl were used to load on gels.

### 4.2 SDS PAGE

Polyacrylamide gel electrophoresis separates proteins in an electrical field according to their molecular weight by sodium dodecylsulfat (SDS) which linearizes proteins and gives the proteins a negative charge. The negatively charged proteins can now migrate with different speed through the electrical field only depending on their molecular weight. The higher the molecular weight, the slower proteins migrate. Gels in this work were prepared using the single gel system where only one polyacrylamide gel without stacking layer is used. This method gives a resolution of proteins comparable to the conventional two-gel Laemmli's standard method, but only one gel is used with a rapidly occurring polymerization <sup>160</sup>. 6 to 8% gels were used for this work and had the standard thickness of 0.75 mm. For electrophoresis, the mini protean system III <sup>™</sup> from Bio-Rad was used. Lysates were loaded onto a gel together with 4 µl of PageRuler Prestained Protein Ladder (Thermo Scientific) and run at 80 V for 30 to 120 min depending on the molecular weight of the proteins to be analyzed. The gel was further used for blotting or coomassie staining.

### 4.3 COOMASSIE STAINING

By staining the single gel with coomassie, the by SDS PAGE separated proteins in the samples are made visible. The gels were stained for 20 min with coomassie staining solution and afterwards destained until the protein bands could easily be detected.

## 4.4 WESTERN BLOTTING

After SDS PAGE, the separated proteins are transferred from the gel onto the carrier structure, a PVDF membrane. An electrical field runs rectangular to the gel and the negatively charged proteins are migrating through the field to the membrane. The pattern of separated proteins stays the same by the transfer. When the proteins reach the surface of the PVDF membrane, it comes to hydrophobic and polar interactions which bind the proteins to the membrane.

The set-up of the blot was as follows: two thick and two thin layers of Whatman gel blot paper were soaked in western transfer buffer and put on the cathode half of the blotting apparatus. The PVDF membrane was first activated in methanol, then put in western transfer buffer and positioned on top of the Whatman gel blot papers. Next, the gel was put onto the membrane. Last, two thin and two thick layers of Whatman gel blot papers were put on top of the gel and the anode half of the blotting apparatus was set onto the whole blot. The blotting was performed for 75 min at the current density of 0.8 mA/cm<sup>2</sup>.

## 4.5 DETECTION OF PROTEINS

After blotting, the membrane was dried for 20 min at 37°C and then reactivated in methanol. To exclude nonspecific antibody binding, the PVDF membrane was blocked for 2 h in 5 ml TBS/3% BSA at room temperature.

The primary antibodies were diluted from 1:1 000 to 1:10 000 in TBS/3%BSA according to the specific antibody information and incubated for 1 h at room temperature. The blot was washed afterwards 3 times with 10 ml TBST and then incubated with the secondary antibody (1:2 500 to 1:10 000) for 1 h at room temperature. After 3 further washing steps with 10 ml TBST, the blot was developed either colorimetric or by using chemiluminescence. In the colorimetric development, alkaline phosphatase (AP) catalyzes the generation of a colored product of 5-bromo-4-chloro-3-indolyl phosphate (BCIP) and nitroblue tetrazolium (NBT) that doesn't fade on the blot. 5 ml of a colorimetric developing solution were applied to the blot and incubated till the bands were seen properly on the blot. The reaction was stopped with H<sub>2</sub>O bidest. More sensitive was the chemiluminescence development with horse radish peroxidase (HRP). Light emission is generated as a by-product of the activity of the HRP and can be detected as dark bands on x-ray films. For this, the Immobilon Western HRP Substrate (Merck Millipore) was used. 1 ml of the luminol reagent and 1 ml peroxide solution were pipetted together, incubated for 5 min on the blot and the chemiluminescence was detected using x-ray films.

## 4.6 PVDF MEMBRANE STRIPPING

The PVDF membrane was stripped of all former used antibodies, but not the separated proteins, to use further antibodies on the same blot. The membrane was first washed with 5 ml H<sub>2</sub>O bidest, then with 5 ml stripping solution for 40 min. After another washing step with 5 ml H<sub>2</sub>O bidest, the blot was blocked again as described in 4.5.

---

# Appendix

---

## 1 Supplementary data

### 1.1 STRAIN PMSS1

#### 1.1.1 SEQUENCING OF THE *SABA* AND *SABB* LOCUS (*HOPQI* GENE)

ATGAAAAACGATTTTTACTTTCTCTATCCCTTGCATCGTCATTACTTTATGCTGAAGACAACGGCTTT  
TTTTGTGAGCGCGGGCTATCAAATCGGCGAAGCGGTGCAAAAAGTGAAAAACGCCGACAAGGTGCAAAA  
ACTTTCAGACGCTTATGAACAATTAAGCCGGCTTTTAACTAACGATAGTGGCACAACTCAAAGACAA  
GCGCACAAAGCGATCAACCAAGCGGTTAATAATTTGAACGAACGCACAAAACTTTAGCCGGTGGGACA  
ACCAATCCCCTGCCTATCAAGCCACGCTTTTAGCGTTGAGATCGGTGTTAGGGCTATGGAATAGCATG  
GGTTATGCAGTCATATGCGGAGGTTATACCAAAGTCCAGGCGAAAACAATCAAAAAATTTCCACTA  
CACCGATGAGAATGGCAACGGCACTACAATCAATTGCGGTGGGAGCACAAATAGTAATGGCACTCATA  
GTTCTAGTGGCACAAATACATTTAAAAGCAGACAAAAATGTTTCTCTATCTATTGACCAATACGAAGCC  
ATCCATGAAGCCTATCAAATCCTTTCAAAGCTTTAAAACAAGCCGGGCTTGCTCCTTTAAATAGCAA  
AGGGGAAAAATTAGAAGCGCATGTAACCACATCAAAGTATCAATCAGATAATCAACTAAAACGACAA  
CTTCTGTTATTGATACGACTAATGATGCGCAAAATCTTTTGACTCAAGCGCAAACGATTGTCAATACC  
CTTAAAGATTATTGCCCATGTTGATAGCGAAATCTAGTAGTGGAAAGTAGTGGCGGAGCTGCTACAAA  
CACCCCTTCATGGCAAACAGCCGGTGGCGGCAAGAATTCATGCGCGACTTTTGGTGGCGGAGTTTAGTGC  
CGCTTCAGACATGATTAATAATGCGCAAAAAATCGTTCAAGAAACCCAACAACCTCAGCGCCAACCAAC  
CAAAAAATATCACACAACCCATAATCTCAACCTTAACACCCCTAGCAGTCTTACGGCTTTAGCTCAAA  
AAATGCTCAAAAACGCGCAATCTCAAGCAGAAATTTTAAAACCTAGCCAATCAAGTGGAGAGCGATTTT  
AACAACTTTCTTCAGGCCATCTTAAAGACTACATAGGGAAATGCGATGCGAGCGCTATAAGCAGCGC  
GAATATGACAATGCAAATCAAAGAACAATTGGGGGAATGGGTGTGCTAGCGTGGAAAGAACTCAG  
TCTTTGTTAAAAACAAGCGCCGCTGATTTTAAACAACCAACGCCTCAAATCAATCAAGCGCAAACCT  
AGCCAACACCCTTATTCAAGAACTTGGCAACAACCTTTTAGGAATATGGGCATGATCGCTTCTTCAAC  
CACGAATAACGGGGCGATGAATGGCCTTGGGGTGCAAGTGGGTTATAACAATTTTTTGGGGAAAAGA  
AAAGATGGGGGTTAAGGTATTATGGTTTCTTTGATTACAACCACGCCTACATCAAATCAAGCTTCTTT  
AATTCGGCTTCTGATGTGTGGACTTATGGGGTGGGCAGCGATTTATTGTTAATTTTCATCAATGATAA  
AAACACCAACTTTTTAGGTAAAAATAACCAGATTTCTTTTGGGCTTTTTGGAGGCATCGCCTTAGCAG  
GGACTTCATGGCTTAATTTCTCAATTCGTGAATTTAAAACCATCAGCAATGTCTATAGCGCTAAAGTG  
AATACCGCTAACTTCCAATTTTTATTCAATTTGGGCTTGAGAACCAATCTCGCTAGACCTAAGAAAAA

AGATAGCCATCATGCGGCTCAACATGGCATGGAATTGGGCGTGAAAATCCCTACCATTAACACGAATT  
ACTATTCTTTTCTAGACACTAACTAGAATATAGGAGGCTTTATAGCGTGTATCTCAATTATGTGTTT  
GCTTACTGA

### 1.1.2 SEQUENCING OF THE *HOPQ* LOCUS (*HOPQII* GENE)

ATGAAAAAACGAAAAAACGATTCTACTTTCTCTAACTCTTGCGTCATCATTGCTCCATGCTGAAGA  
CAACGGCGTTTTTTTAAGCGTGGGCTATCAAATCGGTGAAGCGGTTCAAAAAGTGAAAAACGCCGACA  
AGGTGCAAAAACCTTTCAGACGCTTATGAAAACCTTGAACAAGATTTTAGCTAATCATGACCACTCCAAT  
CCAGAAGCGATTAATACAAACAGCGCCACAGCGATCAATCAAGCGATTGGTAATTTAAACGCAAACAC  
GCAAAATTTAATTGATAAAACAGACAATTCCCCTGCCTATCAAGCCACGCTTTTAGCGCTAAAATCCAC  
GGTGGGGTTATGGAATAGCATAGCTTATGCCGTCATATGCGGAGGCTATACGGATAAACCCAACCACA  
ACATCACAGAACTTTTTACAACCAGCCAGGACAAAATTCAAATTCGATTACTTGCGGTAGTAATGGT  
TTAGGGACTCTTCCAGCGGGCAAAAATTCTCATCTGTCCATTGAACAATTTGCAACGCTCAACAAAGCG  
TATCAAATTATCCAAGCCGCTTTGAAACAAGGTCTCCCTGCTTTAAGCGATACAAAAAACGGTGGA  
AGTAACCATTA AACAGCAACCAACGCTCAAAACATTAATGTCAATAATAACAACAACAATGCTGCTG  
ATGCTACAATTGAAACAAAGAATACTTATATTAACGATGCGCAAAATCTTTTAACCCAAGCGCAAACC  
ATCATCAACACCCTTCAAGACAATTGCCCGATGTTGAAAGGGAAGTCTAGTAGTGGAACATAATGGCGC  
AAACACCCTTCATGGCAAACAAGCGCTAACCAAAATTCGTGCAGCGTTTTTTGGCACGGAATTTAGCGC  
TATTTTCAGACATGATCAGTAACGCTCAAAACATCGTTCAAGAAACCCAACAGCTTAATACTACCCCACT  
AAAAAGCATCGCACAAACCCACAATTTCAACCTTAACTCCCCTAATAGTGTCGCTTTGGCTCAAAGCAT  
GCTCAAAAACGCTCAATCTCAAGCAGCGGTTTTAAACTAGCCAATCAAGTGGGGAACGATTTCAATA  
GAATTTCTACAGGAGTTCTTAAAAATTATATAGAAGAATGCAATGCGAATGCTTCAAGTGAAAGCGTT  
TCTAATAACACTTGGGGGAAAGGTTGCGCGGGCGTGAAACAACTCTAACTTCGCTAGAAAGTAGCAA  
CGCTTCTTTTTCTAGCCAAACGCCTCAAATCAATCAAGCTGAAACCATCGCTAACACTATTGTTCAAGA  
ACTCGGTCATAACCTTTCAAACGGGTGGGCATCATTAGCTCTCAAACCAATAACGGGGCGATGAACG  
GCCTTGGGGTGCAAGTGGGTTATAACAATTTTTTGGGGAAAAGAAAAGATGGGGGTTAAGGTATTA  
TGGTTTCTTTGATTACAACCACGCCTACATCAAATCAAGCTTCTTTAATTCGGCTTCTGATGTGTGGAC  
TTATGGGGTGGCAGCGATTTATTGTTTAATTTTCATCAATGATAAAAAACCAACTTTTTAGGTAAAA  
ATAACCAGATTTCTTTTGGGCTTTTTGGAGGCATCGCCTTAGCAGGGACTTCATGGCTTAATTTCTCAAT  
TCGTGAATTTAAAAACCATCAGCAATGTCTATAGCGCTAAAGTGAATACCGCTAACTTCCAATTTTTTA  
TTCAATTTGGGCTTGAGAACCAATCTCGCTAGACCTAAGAAAAAGATAGCCATCATGCGGCTCAACA  
TGGCATGGAATTTGGCGTGAAAATCCCTACCATTAACACGAATTATTATTCTTTTCTAGACACTAAAC  
TAGAATATAGGAGGCTTTATAGCGTGTATCTCAATTATGTGTTTGCCTATTA

## 1.2 STRAIN SS1

### 1.2.1 SEQUENCING OF THE *SAB*A AND *SAB*B LOCUS (*HOP*QI GENE)

ATGAAAAACGATTTTTACTTTCTCTATCCCTTGCATCGTCATTACTTTATGCTGAAGACAACGGCTTT  
TTTGTGAGCGCGGGCTATCAAATCGGCGAAGCGGTGCAAAAAGTGAAAAACGCCGACAAGGTGCAAAA  
ACTTTCAGACGCTTATGAACAATTAAGCCGGCTTTTAACTAACGATAGTGGCACAACCTCAAAGACAA  
GCGCACAAGCGATCAACCAAGCGGTTAATAATTTGAACGAACGCACAAAACTTTAGCCGGTGGGACA  
ACCAATCCCCCTGCCTATCAAGCCACGCTTTTAGCGTTGAGATCGGTGTTAGGGCTATGGAATAGCATG  
GGTTATGCAGTCATATGCGGAGGTTATACCAAAGTCCAGGCGAAAAACAATCAAAAAAATTTCCACTA  
CACCGATGAGAATGGCAACGGCACTACAATCAATTGCGGTGGGAGCACAATAGTAATGGCACTCATA  
GTTCTAGTGGCACAATAACATTAAGCAGACAAAAATGTTTCTCTATCTATTGACCAATACGAAGCC  
ATCCATGAAGCCTATCAAATCCTTTCAAAGCTTTAAAACAAGCCGGGCTTGCTCCTTTAAATAGCAA  
AGGGGAAAAATTAGAAGCGCATGTAACCACATCAAAGTATCAATCAGATAATCAAACCTAAAACGACAA  
CTTCTGTTATTGATACGACTAATGATGCGCAAAATCTTTTACTCAAGCGCAAACGATTGTCAATACC  
CTTAAAGATTATTGCCCATGTTGATAGCGAAATCTAGTAGTGAAGTAGTGGCGGAGCTGCTACAAA  
CACCCCTTCATGGCAAACAGCCGGTGGCGCAAGAATTCATGCGCGACTTTTGGTGGGAGTTTAGTGC  
CGCTTCAGACATGATTAATAATGCGCAAAAAATCGTTCAAGAAACCCAACAACCTCAGCGCCAACCAAC  
CAAAAAATATCACACAACCCATAATCTCAACCTTAACACCCCTAGCAGTCTTACGGCTTTAGCTCAAA  
AAATGCTCAAAAACGCGCAATCTCAAGCAGAAATTTTAAAACCTAGCCAATCAAGTGGAGAGCGATTTT  
AACAACTTTCTTCAGGCCATCTTAAAGACTACATAGGGAAATGCGATGCGAGCGCTATAAGCAGCGC  
GAATATGACAATGCAAAATCAAAGAACAATTGGGGGAATGGGTGTGCTAGCGTGAAGAACTCAG  
TCTTTGTTAAAAACAAGCGCCGCTGATTTTAAACAACCAACGCCTCAAATCAATCAAGCGCAAAACCT  
AGCCAACACCCTTATTCAAGAACTTGGCAACAACCCTTTTAGGAATATGGGCATGATCGCTTCTTCAAC  
CACGAATAACGGGGCGATGAATGGCCTTGGGGTGCAAGTGGGTTATAACAATTTTTTGGGGAAAAGA  
AAAGATGGGGGTTAAGGTATTATGGTTCTTTGATTACAACCACGCCTACATCAAATCAAGCTTCTTT  
AATTCGGCTTCTGATGTGTGGACTTATGGGGTGGGCAGCGATTTATTGTTAATTTTCATCAATGATAA  
AAACACCAACTTTTTAGGTAAAAATAACCAGATTTCTTTTGGGCTTTTTGGAGGCATCGCCTTAGCAG  
GGACTTCATGGCTTAATTCTCAATTCGTGAATTTAAAACCATCAGCAATGTCTATAGCGCTAAAGTG  
AATACCGCTAACTTCCAATTTTTATTCAATTTGGGCTTGAGAACCAATCTCGCTAGACCTAAGAAAA  
AGATAGCCATCATGCGGCTCAACATGGCATGGAATTTGGGCGTAAAATCCCTACCATTAACACGAATT  
ACTATTCTTTTCTAGACACTAACTAGAATATAGGAGGCTTTATAGCGTGTATCTCAATTATGTGTTT  
GCTTACTGA

### 1.2.2 SEQUENCING OF THE *HOPQ* LOCUS (*HOPQII* GENE)

ATGAAAAAACGAAAAACGATTCTACTTTCTCTAACTCTTGCGTCATCATTGCTCCATGCTGAAGA  
CAACGGCGTTTTTTAAGCGTGGGCTATCAAATCGGTGAAGCGGTTCAAAAAGTGAAAAACGCCGACA  
AGGTGCAAAAACCTTTCAGACGCTTATGAAAACCTTGAACAAGATTTTAGCTAATCATGACCACTCCAAT  
CCAGAAGCGATTAATACAAACAGCGCCACAGCGATCAATCAAGCGATTGGTAATTTAAACGCAAACAC  
GCAAAATTTAATTGATAAAACAGACAATTCCCCTGCCTATCAAGCCACGCTTTTAGCGCTAAAATCCAC  
GGTGGGGTTATGGAATAGCATAGCTTATGCCGTCATATGCCGAGGCTATACGGATAAACCCAACCACA  
ACATCACAGAACTTTTTACAACCAGCCAGGACAAAATTCAAATTCGATTACTTGGCGTAGTAATGGT  
TTAGGGACTCTTCCAGCGGGCAAAAATTCATCTGTCCATTGAACAATTTGCAACGCTCAACAAAGCG  
TATCAAATTTATCCAAGCCGCTTTGAAACAAGGTCTCCCTGCTTTAAGCGATACAAAAAACGGTGGA  
AGTAACCATTAACAGCAACCAACGCTCAAAACATTAATGTCAATAATAACAACAACAATGCTGCTG  
ATGCTACAATTGAAACAAGAATACTTATATTAACGATGCGCAAAATCTTTTAACCCAAGCGCAAACC  
ATCATCAACACCCTTCAAGACAATTGCCCGATGTTGAAAGGGAAGTCTAGTAGTGGAACATAATGGCGC  
AAACACCCTTCATGGCAAACAAGCGCTAACCAAAATTCGTGCAGCGTTTTTTGGCACGGAATTTAGCGC  
TATTTTCAGACATGATCAGTAACGCTCAAAACATCGTTCAAGAAACCCAACAGCTTAATACTACCCCACT  
AAAAAGCATCGCACAACCCCACAATTTCAACCTTAACTCCCCTAATAGTGTCGCTTTGGCTCAAAGCAT  
GCTCAAAAACGCTCAATCTCAAGCAGCGGTTTTAAACTAGCCAATCAAGTGGGGAACGATTTCAATA  
GAATTTCTACAGGAGTTCTTAAAAATTATATAGAAGAATGCAATGCGAATGCTTCAAGTGAAAGCGTT  
TCTAATAACACTTGGGGGAAAGGTTGCGCGGGCGTGAAACAACTCTAACTTCGCTAGAAAGTAGCAA  
CGCTTCTTTTTCTAGCCAAACGCCTCAAATCAATCAAGCTGAAACCATCGCTAACACTATTGTTCAAGA  
ACTCGGTCATAACCCCTTCAAACGGGTGGGCATCATTAGCTCTCAAACCAATAACGGGGCGATGAACG  
GCCTTGGGGTGCAAGTGGGTTATAAACAATTTTTTGGGGAAAAGAAAAGATGGGGGTTAAGGTATTA  
TGGTTTCTTTGATTACAACCACGCCTACATCAAATCAAGCTTCTTTAATTCGGCTTCTGATGTGTGGAC  
TTATGGGGTGGGCAGCGATTTATTGTTTAATTCATCAATGATAAAAAACCAACTTTTTTAGGTAAAA  
ATAACCAGATTTCTTTTGGGCTTTTTGGAGGCATCGCTTAGCAGGGACTTCATGGCTTAATCTCAAT  
TCGTGAATTTAAAAACCATCAGCAATGTCTATAGCGCTAAAGTGAATACCGCTAACTTCCAATTTTTA  
TTCAATTTGGGCTTGAGAACCAATCTCGCTAGACCTAAGAAAAAGATAGCCATCATGCGGCTCAACA  
TGGCATGGAATTTGGCGTGAAAATCCCTACCATTAACACGAATTATTATCTTTTCTAGACACTAAAC  
TAGAATATAGGAGGCTTTATAGCGTGTATCTCAATTATGTGTTTGCCTATTAA

### 1.3 PROTEIN SEQUENCE HOPQI (PMSS1/SS1)

MKKRFLLSLASSLLYAEDNGFFVSAGYQIGEAVQKVKNADKVQKLSDAYEQLSRLLTNDSGTNSKTS  
 QAINQAVNNLNERTKTLAGGTTNSPAYQATLLALRSVLGLWNSMGYAVICGGYTKSPGENNQKNFHYT  
 DENGNGTTINCGGSTNSNGTHSSSGTNTLKADKNVSLSIDQYEAHEAYQILSKALKQAGLAPLNSKGEK  
 LEAHVTTSKYQSDNQTKTTTTVIDTTNDAQNLLTQAQTIVNTLKDYCPMLIAKSSSGSSGGAATNTPSW  
 QTAGGGKNSCATFGAEFSAASDMINNAQKIVQETQQLSANQPKNITQPHNLNLNTPSSLTALAQKMLK  
 NAQSQAIEILKLANQVESDFNKLSSGHLKDYIGKCDASAISSANMTMQNQKNNWGNGCASVEETQSLK  
 TSAADFNNQTPQINQAQNLANTLIQELGNNPFRNMGMIASTTNNGAMNGLGVQVGYKQFFGEKKRW  
 GLRYYGFFDYNHAYIKSSFFNSASDVWPTYGVGSDLLFNFINDKNTNFLGKNNQISFGLFGGIALAGTSWL  
 NSQFVNLTISNVYSAKVNTANFQFLFNGLRNTLARPKKKDSHHAHQHGMELGVKIPTINTNYYSFLD  
 TKLEYRRLYSVYLNYYVFAY

### 1.4 PROTEIN SEQUENCE HOPQII (PMSS1/SS1)

MKKTKKTILLSLTLASSLLHAEDNGVFLSVGYQIGEAVQKVKNADKVQKLSDAYENLNKILANHDHNSP  
 EAINTNSATAINQAIGNLNANTQNLIDKTDNSPAYQATLLALKSTVGLWNSIAYAVICGGYTDKPNHNIT  
 ETFYNQPGQNSNSITCGSNGLGTLGKNSHLSIEQFATLNKAYQIIQAALKQGLPALSDTKKTVEVTIKT  
 ATNAQNINVNNNNNAADATIETKNYINDAQNLLTQAQTIINTLQDNCPLKGGKSSSGTNGANTPSW  
 QTSANQNSCSVFGTEFSAISDMISNAQNIVQETQQLNTTPLKSIAQPHNFNLNSPNSVALAQSMKNAQS  
 QAAVLKLANQVGNDFNRISTGVLKNYIEECNANASSESVSNNWTWGKGCAGVKQTLTSLESSNASFSSQT  
 PQINQAETIANTIVQELGHNPFKRVGISSQTNNGAMNGLGVQVGYKQFFGEKKRWGLRYYGFFDYNHA  
 YIKSSFFNSASDVWPTYGVGSDLLFNFINDKNTNFLGKNNQISFGLFGGIALAGTSWLNQFVNLTISNV  
 YSAKVNTANFQFLFNGLRNTLARPKKKDSHHAHQHGMELGVKIPTINTNYYSFLD  
 TKLEYRRLYSVYLNYYVFAY

## 2 List of figures

FIGURE 1: TYPICAL IG DOMAIN OF TWO $\beta$ -SHEETS. ....	5
FIGURE 2: CEACACM FAMILY MEMBERS. ....	8
FIGURE 3: CEACAM1 ISOFORMS BY ALTERNATIVE SPLICING. ....	10
FIGURE 4: PSGs. ....	14
FIGURE 5: CGF MEMBERS IN MAMMALS. ....	17
FIGURE 6: BACTERIAL SPECIES EXAMINED TO BIND CEACAM RECEPTORS. ....	20
FIGURE 7: MAJOR VIRULENCE FACTORS OF <i>H. PYLORI</i> . ....	24
FIGURE 8: HOP AND HOR FAMILY TREE OF OMPs. ....	26
FIGURE 9: HUMAN CEACAMs 1, 3, 5 AND 6. ....	31
FIGURE 10: CONTROLS FOR THE BACTERIAL PULL-DOWN ASSAY. ....	32
FIGURE 11: <i>H. PYLORI</i> STRAIN P12 IS BINDING TO CEA1-N-GFP AND CEA5-N-GFP. ....	34
FIGURE 12: TEST OF FURTHER <i>H. PYLORI</i> STRAINS FOR THEIR CAPABILITY TO BIND CEACAM1 AND CEACAM5. ....	36
FIGURE 13: EXCLUSIVE AND SPECIFIC BINDING OF THE GENUS <i>H. PYLORI</i> TO HUMAN CEACAMs. ....	39
FIGURE 14: GENERATION OF MARKERFREE OUTER MEMBRANE PROTEIN MUTANTS OF <i>H. PYLORI</i> . ....	43
FIGURE 15: HOPQ IS THE ADHESIN INTERACTING WITH CEACAM1 AND CEACAM5. ....	45
FIGURE 16: CLUSTALW ALIGNMENT OF HOPQI AND HOPQII. ....	48
FIGURE 17: GENERATION AND CHARACTERIZATION OF THE ANTIBODY AK298 AGAINST HOPQI. ....	51
FIGURE 18: SYNTHESIS OF CEACAM1, 5 AND 6 ON EPITHELIAL CELL LINES. ....	53
FIGURE 19: HEK293 CELLS PRODUCING HUMAN CEACAM1 AND CEACAM5, RESPECTIVELY. ....	56
FIGURE 20: <i>IN VITRO</i> INFECTION OF HEK293/CEACAM5 AND AGS CELLS. ....	58
FIGURE 21: KNOCKDOWN OF CEACAM1, CEACAM5 AND CEACAM6 IN AGS CELLS. ....	59
FIGURE 22: EXCERPT FROM KUESPERT <i>ET AL.</i> , 2007 AND ROTH <i>ET AL.</i> , 2013. ....	62

### 3 List of tables

TABLE 1: LIST OF CELL LINES. ....	81
TABLE 2: <i>E. COLI</i> THAT WERE USED IN THIS WORK. ....	81
TABLE 3: <i>H. PYLORI</i> STRAINS THAT WERE USED IN THIS WORK. ....	82
TABLE 4: <i>M. CATARRHALIS</i> AND <i>N. GONORRHOEAE</i> . ....	82
TABLE 5: LIST OF PLASMIDS GENERATED OR USED IN THIS WORK. ....	83
TABLE 6: LIST OF OLIGONUCLEOTIDES. ....	85
TABLE 7: LIST OF SIRNA. ....	86
TABLE 8: LIST OF COMMERCIALY AVAILABLE KITS. ....	86
TABLE 9: LIST OF ANTIBODIES. ....	87
TABLE 10: LIST OF CULTURE MEDIUM FOR BACTERIAL GROWTH. ....	88
TABLE 11: LIST OF BUFFERS AND SOLUTIONS. ....	89
TABLE 12: STANDARD REACTION CONDITIONS TAQ POLYMERASE (PAN BIOTECH GMBH). ....	100
TABLE 13: STANDARD REACTION CONDITIONS EX TAQ DNA POLYMERASE (TAKARA). ....	100
TABLE 14: STANDARD PCR PROGRAM USED FOR THE DNA POLYMERASES (TAKARA/PAN). ....	101
TABLE 15: STANDARD REACTION CONDITIONS RAPD PCR. ....	102
TABLE 16: RAPD PCR PROGRAM. ....	102

## 4 Abbreviations

aa	amino acid
AOM	azoxymethane
BabA	blood group antigen binding protein A
BB	Brucella broth
c-Abl	<i>Abelson murine leukemia viral oncogene homolog</i> , tyrosine-protein kinase
CagA	cytotoxin associated gene A
cagPAI	cag pathogenicity island
CAM	cell adhesion molecules
cat	chloramphenicol
CD	cluster of differentiation
CEACAM	carcinoembryonic antigen related cell adhesion molecules
CEA-N	N-terminal domain of CEACAM5
CGF	CEA gene family
CS	chondroitin sulfate
Da	Dalton
ECM	extracellular matrix
erm	erythromycin
FCS	fetal calf serum
FN	fibronectin
GAGs	glycosaminoglycans
GPI-anchor	Glycosylphosphatidylinositol-anchor
hnRNP M	heterogeneous nuclear ribonucleoprotein M
Hop	<i>Helicobacter pylori</i> outer membrane protein
Hor	<i>Helicobacter pylori</i> outer membrane protein related
HS	heparan sulfate
IgC like domain	Immunoglobulin like constant domain
IgSF	Immunoglobulin superfamily
IgV like domain	Immunoglobulin like variable domain
ITAM	immunoreceptor tyrosine-based activation motif
ITIM	immunoreceptor tyrosine-based inhibitory motif
LabA	lacdiNAc-binding adhesin
lacdiNAc	N,N'-diacetyllactosediamine
LB	Luria Bertani
MBP	Maltose binding protein
MFI	mean fluorescence intensity
MHC	major histocompatibility complex
MOI	multiplicity of infection
MW	molecular weight
NCA	non-specific cross reacting antigen
OD	optical density
OipA	outer inflammatory protein A
OMP	outer membrane protein
Op <sub>CEA</sub>	opacity-associated proteins binding to CEACAMs
PAGE	polyacrylamide gel electrophoresis
PAI	pathogenicity island

---

PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
PI3K	phosphatidylinositol 3'-kinase
PIP	phosphatidylinositol phosphate
PSG	pregnancy specific glycoproteins
P-Tyr	tyrosine phosphorylation
PVDF	Polyvinylidenfluorid
RAPD	randomly amplified polymorphic DNA
RNAi	RNA interference
rpsL	ribosomal protein S12
SabA	sialic acid binding adhesin A
sCEACAM1	secreted forms of CEACAM1
SDS	sodium dodecylsulfat
SH2	SRC homology 2-domain
SHP1	SH2-domain containing protein tyrosine phosphatase 1
src kinase	( <i>sarcoma</i> ), cellular tyrosine-protein kinase
SSM	slipped strand mispairing
T4SS	type IV secretion system
TAE-buffer	Tris-acetate-EDTA buffer
TCR	T-cell receptor
TGF $\beta$ 1	transforming growth factor beta-1
TIM-3	T-cell immunoglobulin domain and mucin domain-3
TLR	toll-like receptor
UspA1	ubiquitous surface protein A1
VacA	vacuolating cytotoxin A

---

# Literature

---

1. Barclay AN. Membrane proteins with immunoglobulin-like domains--a master superfamily of interaction molecules. *Seminars in immunology* 2003, **15**(4): 215-223.
2. Bork P, Holm L, Sander C. The immunoglobulin fold. Structural classification, sequence patterns and common core. *J Mol Biol* 1994, **242**(4): 309-320.
3. Williams AF, Barclay AN. The immunoglobulin superfamily--domains for cell surface recognition. *Annual review of immunology* 1988, **6**: 381-405.
4. Smith DK, Xue H. Sequence profiles of immunoglobulin and immunoglobulin-like domains. *J Mol Biol* 1997, **274**(4): 530-545.
5. Harpaz Y, Chothia C. Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains. *J Mol Biol* 1994, **238**(4): 528-539.
6. Shimono Y, Rikitake Y, Mandai K, Mori M, Takai Y. Immunoglobulin superfamily receptors and adherens junctions. *Sub-cellular biochemistry* 2012, **60**: 137-170.
7. Brummendorf T, Rathjen FG. Structure/function relationships of axon-associated adhesion receptors of the immunoglobulin superfamily. *Current opinion in neurobiology* 1996, **6**(5): 584-593.
8. van den Berg TK, Yoder JA, Litman GW. On the origins of adaptive immunity: innate immune receptors join the tale. *Trends in immunology* 2004, **25**(1): 11-16.
9. Marchalonis JJ. Lymphocyte surface immunoglobulins. *Science* 1975, **190**(4209): 20-29.
10. Korman AJ, Auffray C, Schamboeck A, Strominger JL. The amino acid sequence and gene organization of the heavy chain of the HLA-DR antigen: homology to immunoglobulins. *Proc Natl Acad Sci U S A* 1982, **79**(19): 6013-6017.
11. Olive D. [Lymphocyte coreceptors]. *Medecine sciences : M/S* 2006, **22**(12): 1069-1074.
12. Linke WA. Stretching molecular springs: elasticity of titin filaments in vertebrate striated muscle. *Histology and histopathology* 2000, **15**(3): 799-811.
13. Siegel GJ, Agranoff BW, Albers RW, editors. *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*. 1999, **6th edition**.

14. Hammarstrom S. The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. *Seminars in cancer biology* 1999, **9**(2): 67-81.
15. Teglund S, Olsen A, Khan WN, Frangsmyr L, Hammarstrom S. The pregnancy-specific glycoprotein (PSG) gene cluster on human chromosome 19: fine structure of the 11 PSG genes and identification of 6 new genes forming a third subgroup within the carcinoembryonic antigen (CEA) family. *Genomics* 1994, **23**(3): 669-684.
16. Brummendorf T, Rathjen FG. Cell adhesion molecules 1: immunoglobulin superfamily. *Protein profile* 1995, **2**(9): 963-1108.
17. Volkmer H, Schreiber J, Rathjen FG. Regulation of adhesion by flexible ectodomains of IgCAMs. *Neurochemical research* 2013, **38**(6): 1092-1099.
18. Paxton RJ, Mooser G, Pande H, Lee TD, Shively JE. Sequence analysis of carcinoembryonic antigen: identification of glycosylation sites and homology with the immunoglobulin supergene family. *Proc Natl Acad Sci U S A* 1987, **84**(4): 920-924.
19. Rudd PM, Wormald MR, Stanfield RL, Huang M, Mattsson N, Speir JA, *et al.* Roles for glycosylation of cell surface receptors involved in cellular immune recognition. *J Mol Biol* 1999, **293**(2): 351-366.
20. Pavlopoulou A, Scorilas A. A comprehensive phylogenetic and structural analysis of the carcinoembryonic antigen (CEA) gene family. *Genome biology and evolution* 2014, **6**(6): 1314-1326.
21. Czepczynska-Krezel H, Krop-Watorek A. [Human carcinoembryonic antigen family proteins, structure and function]. *Postepy higieny i medycyny doswiadczalnej* 2012, **66**: 521-533.
22. Skubitz KM, Skubitz AP. Interdependency of CEACAM-1, -3, -6, and -8 induced human neutrophil adhesion to endothelial cells. *Journal of translational medicine* 2008, **6**: 78.
23. Beauchemin N, Draber P, Dveksler G, Gold P, Gray-Owen S, Grunert F, *et al.* Redefined nomenclature for members of the carcinoembryonic antigen family. *Experimental cell research* 1999, **252**(2): 243-249.
24. Prall F, Nollau P, Neumaier M, Haubeck HD, Drzeniek Z, Helmchen U, *et al.* CD66a (BGP), an adhesion molecule of the carcinoembryonic antigen family, is expressed in epithelium, endothelium, and myeloid cells in a wide range of normal human tissues. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* 1996, **44**(1): 35-41.

25. Chen T, Grunert F, Medina-Marino A, Gotschlich EC. Several carcinoembryonic antigens (CD66) serve as receptors for gonococcal opacity proteins. *The Journal of experimental medicine* 1997, **185**(9): 1557-1564.
26. Ocklind C, Obrink B. Intercellular adhesion of rat hepatocytes. Identification of a cell surface glycoprotein involved in the initial adhesion process. *The Journal of biological chemistry* 1982, **257**(12): 6788-6795.
27. Gray-Owen SD, Blumberg RS. CEACAM1: contact-dependent control of immunity. *Nature reviews Immunology* 2006, **6**(6): 433-446.
28. Edlund M, Blikstad I, Obrink B. Calmodulin binds to specific sequences in the cytoplasmic domain of C-CAM and down-regulates C-CAM self-association. *The Journal of biological chemistry* 1996, **271**(3): 1393-1399.
29. Schumann D, Chen CJ, Kaplan B, Shively JE. Carcinoembryonic antigen cell adhesion molecule 1 directly associates with cytoskeleton proteins actin and tropomyosin. *The Journal of biological chemistry* 2001, **276**(50): 47421-47433.
30. Nagaishi T, Iijima H, Nakajima A, Chen D, Blumberg RS. Role of CEACAM1 as a regulator of T cells. *Annals of the New York Academy of Sciences* 2006, **1072**: 155-175.
31. Shi JF, Xu SX, He P, Xi ZH. Expression of carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) and its correlation with angiogenesis in gastric cancer. *Pathology, research and practice* 2014, **210**(8): 473-476.
32. Greicius G, Severinson E, Beauchemin N, Obrink B, Singer BB. CEACAM1 is a potent regulator of B cell receptor complex-induced activation. *Journal of leukocyte biology* 2003, **74**(1): 126-134.
33. Hosomi S, Chen Z, Baker K, Chen L, Huang YH, Olszak T, *et al.* CEACAM1 on activated NK cells inhibits NKG2D-mediated cytolytic function and signaling. *European journal of immunology* 2013, **43**(9): 2473-2483.
34. van Gisbergen KP, Ludwig IS, Geijtenbeek TB, van Kooyk Y. Interactions of DC-SIGN with Mac-1 and CEACAM1 regulate contact between dendritic cells and neutrophils. *FEBS letters* 2005, **579**(27): 6159-6168.
35. Dery KJ, Gaur S, Gencheva M, Yen Y, Shively JE, Gaur RK. Mechanistic control of carcinoembryonic antigen-related cell adhesion molecule-1 (CEACAM1) splice isoforms by the heterogeneous nuclear ribonuclear proteins hnRNP L, hnRNP A1, and hnRNP M. *The Journal of biological chemistry* 2011, **286**(18): 16039-16051.
36. Singer BB, Scheffrahn I, Obrink B. The tumor growth-inhibiting cell adhesion molecule CEACAM1 (C-CAM) is differently expressed in proliferating and quiescent epithelial cells and regulates cell proliferation. *Cancer research* 2000, **60**(5): 1236-1244.

37. Klaile E, Vorontsova O, Sigmundsson K, Muller MM, Singer BB, Ofverstedt LG, *et al.* The CEACAM1 N-terminal Ig domain mediates cis- and trans-binding and is essential for allosteric rearrangements of CEACAM1 microclusters. *J Cell Biol* 2009, **187**(4): 553-567.
38. Obrink B. CEA adhesion molecules: multifunctional proteins with signal-regulatory properties. *Curr Opin Cell Biol* 1997, **9**(5): 616-626.
39. Huang YH, Zhu C, Kondo Y, Anderson AC, Gandhi A, Russell A, *et al.* CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. *Nature* 2014.
40. Brummer J, Ebrahimnejad A, Flayeh R, Schumacher U, Loning T, Bamberger AM, *et al.* cis Interaction of the cell adhesion molecule CEACAM1 with integrin beta(3). *The American journal of pathology* 2001, **159**(2): 537-546.
41. Hunter I, Sawa H, Edlund M, Obrink B. Evidence for regulated dimerization of cell-cell adhesion molecule (C-CAM) in epithelial cells. *The Biochemical journal* 1996, **320 ( Pt 3)**: 847-853.
42. Patel PC, Lee HS, Ming AY, Rath A, Deber CM, Yip CM, *et al.* Inside-out signaling promotes dynamic changes in the carcinoembryonic antigen-related cellular adhesion molecule 1 (CEACAM1) oligomeric state to control its cell adhesion properties. *The Journal of biological chemistry* 2013, **288**(41): 29654-29669.
43. Voges M, Bachmann V, Naujoks J, Kopp K, Hauck CR. Extracellular IgC2 constant domains of CEACAMs mediate PI3K sensitivity during uptake of pathogens. *PLoS One* 2012, **7**(6): e39908.
44. Pils S, Gerrard DT, Meyer A, Hauck CR. CEACAM3: an innate immune receptor directed against human-restricted bacterial pathogens. *International journal of medical microbiology : IJMM* 2008, **298**(7-8): 553-560.
45. Hauck CR, Meyer TF, Lang F, Gulbins E. CD66-mediated phagocytosis of Opa52 *Neisseria gonorrhoeae* requires a Src-like tyrosine kinase- and Rac1-dependent signalling pathway. *EMBO J* 1998, **17**(2): 443-454.
46. McCaw SE, Schneider J, Liao EH, Zimmermann W, Gray-Owen SD. Immunoreceptor tyrosine-based activation motif phosphorylation during engulfment of *Neisseria gonorrhoeae* by the neutrophil-restricted CEACAM3 (CD66d) receptor. *Mol Microbiol* 2003, **49**(3): 623-637.
47. Pils S, Kopp K, Peterson L, Delgado Tascon J, Nyffenegger-Jann NJ, Hauck CR. The adaptor molecule Nck localizes the WAVE complex to promote actin polymerization during CEACAM3-mediated phagocytosis of bacteria. *PLoS One* 2012, **7**(3): e32808.
48. Kopp K, Buntru A, Pils S, Zimmermann T, Frank R, Zumbusch A, *et al.* Grb14 is a negative regulator of CEACAM3-mediated phagocytosis of pathogenic bacteria. *The Journal of biological chemistry* 2012, **287**(46): 39158-39170.

49. Buntru A, Kopp K, Voges M, Frank R, Bachmann V, Hauck CR. Phosphatidylinositol 3'-kinase activity is critical for initiating the oxidative burst and bacterial destruction during CEACAM3-mediated phagocytosis. *The Journal of biological chemistry* 2011, **286**(11): 9555-9566.
50. Kuespert K, Pils S, Hauck CR. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol* 2006, **18**(5): 565-571.
51. Gold P, Freedman SO. Demonstration of Tumor-Specific Antigens in Human Colonic Carcinomata by Immunological Tolerance and Absorption Techniques. *The Journal of experimental medicine* 1965, **121**: 439-462.
52. Rhesus Macaque Genome S, Analysis C, Gibbs RA, Rogers J, Katze MG, Bumgarner R, *et al.* Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 2007, **316**(5822): 222-234.
53. Thompson JA. Molecular cloning and expression of carcinoembryonic antigen gene family members. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* 1995, **16**(1): 10-16.
54. Screatton RA, DeMarte L, Draber P, Stanners CP. The specificity for the differentiation blocking activity of carcinoembryonic antigen resides in its glycoposphatidyl-inositol anchor. *J Cell Biol* 2000, **150**(3): 613-626.
55. Camacho-Leal P, Stanners CP. The human carcinoembryonic antigen (CEA) GPI anchor mediates anoikis inhibition by inactivation of the intrinsic death pathway. *Oncogene* 2008, **27**(11): 1545-1553.
56. Goswami D, Gowrishankar K, Bilgrami S, Ghosh S, Raghupathy R, Chadda R, *et al.* Nanoclusters of GPI-anchored proteins are formed by cortical actin-driven activity. *Cell* 2008, **135**(6): 1085-1097.
57. Thomas P, Forse RA, Bajenova O. Carcinoembryonic antigen (CEA) and its receptor hnRNP M are mediators of metastasis and the inflammatory response in the liver. *Clinical & experimental metastasis* 2011, **28**(8): 923-932.
58. Stern N, Markel G, Arnon TI, Gruda R, Wong H, Gray-Owen SD, *et al.* Carcinoembryonic antigen (CEA) inhibits NK killing via interaction with CEA-related cell adhesion molecule 1. *Journal of immunology* 2005, **174**(11): 6692-6701.
59. Chen J, Li Q, An Y, Lv N, Xue X, Wei J, *et al.* CEACAM6 induces epithelial-mesenchymal transition and mediates invasion and metastasis in pancreatic cancer. *International journal of oncology* 2013, **43**(3): 877-885.
60. Kuroki M, Arakawa F, Matsuo Y, Oikawa S, Misumi Y, Nakazato H, *et al.* Molecular cloning of nonspecific cross-reacting antigens in human granulocytes. *The Journal of biological chemistry* 1991, **266**(18): 11810-11817.

61. Delgado Tascon J, Adrian J, Kopp K, Scholz P, Tschan MP, Kuespert K, *et al.* The granulocyte orphan receptor CEACAM4 is able to trigger phagocytosis of bacteria. *Journal of leukocyte biology* 2015.
62. Singer BB, Opp L, Heinrich A, Schreiber F, Binding-Liermann R, Berrocal-Almanza LC, *et al.* Soluble CEACAM8 interacts with CEACAM1 inhibiting TLR2-triggered immune responses. *PLoS One* 2014, **9**(4): e94106.
63. Zheng J, Miller KK, Yang T, Hildebrand MS, Shearer AE, DeLuca AP, *et al.* Carcinoembryonic antigen-related cell adhesion molecule 16 interacts with alpha-tectorin and is mutated in autosomal dominant hearing loss (DFNA4). *Proc Natl Acad Sci U S A* 2011, **108**(10): 4218-4223.
64. Michaelidou K, Tzovaras A, Missitzis I, Ardavanis A, Scorilas A. The expression of the CEACAM19 gene, a novel member of the CEA family, is associated with breast cancer progression. *International journal of oncology* 2013, **42**(5): 1770-1777.
65. Zhang H, Eisenried A, Zimmermann W, Shively JE. Role of CEACAM1 and CEACAM20 in an in vitro model of prostate morphogenesis. *PLoS One* 2013, **8**(1): e53359.
66. Moore T, Dveksler GS. Pregnancy-specific glycoproteins: complex gene families regulating maternal-fetal interactions. *The International journal of developmental biology* 2014, **58**(2-4): 273-280.
67. Waterhouse R, Ha C, Dveksler GS. Murine CD9 is the receptor for pregnancy-specific glycoprotein 17. *The Journal of experimental medicine* 2002, **195**(2): 277-282.
68. McLellan AS, Zimmermann W, Moore T. Conservation of pregnancy-specific glycoprotein (PSG) N domains following independent expansions of the gene families in rodents and primates. *BMC evolutionary biology* 2005, **5**: 39.
69. Shanley DK, Kiely PA, Golla K, Allen S, Martin K, O'Riordan RT, *et al.* Pregnancy-specific glycoproteins bind integrin alphaIIb beta3 and inhibit the platelet-fibrinogen interaction. *PLoS One* 2013, **8**(2): e57491.
70. Snyder SK, Wessner DH, Wessells JL, Waterhouse RM, Wahl LM, Zimmermann W, *et al.* Pregnancy-specific glycoproteins function as immunomodulators by inducing secretion of IL-10, IL-6 and TGF-beta1 by human monocytes. *American journal of reproductive immunology* 2001, **45**(4): 205-216.
71. Blois SM, Sulkowski G, Tirado-Gonzalez I, Warren J, Freitag N, Klapp BF, *et al.* Pregnancy-specific glycoprotein 1 (PSG1) activates TGF-beta and prevents dextran sodium sulfate (DSS)-induced colitis in mice. *Mucosal Immunol* 2014, **7**(2): 348-358.

72. Chen L, Klass C, Woods A. Syndecan-2 regulates transforming growth factor-beta signaling. *The Journal of biological chemistry* 2004, **279**(16): 15715-15718.
73. Lisboa FA, Warren J, Sulkowski G, Aparicio M, David G, Zudaire E, *et al.* Pregnancy-specific glycoprotein 1 induces endothelial tubulogenesis through interaction with cell surface proteoglycans. *The Journal of biological chemistry* 2011, **286**(9): 7577-7586.
74. Chang CL, Semyonov J, Cheng PJ, Huang SY, Park JI, Tsai HJ, *et al.* Widespread divergence of the CEACAM/PSG genes in vertebrates and humans suggests sensitivity to selection. *PLoS One* 2013, **8**(4): e61701.
75. Kammerer R, Popp T, Hartle S, Singer BB, Zimmermann W. Species-specific evolution of immune receptor tyrosine based activation motif-containing CEACAM1-related immune receptors in the dog. *BMC evolutionary biology* 2007, **7**: 196.
76. Kammerer R, Zimmermann W. Coevolution of activating and inhibitory receptors within mammalian carcinoembryonic antigen families. *BMC biology* 2010, **8**: 12.
77. Naghibalhossaini F, Stanners CP. Minimal mutations are required to effect a radical change in function in CEA family members of the Ig superfamily. *Journal of cell science* 2004, **117**(Pt 5): 761-769.
78. Schmitter T, Agerer F, Peterson L, Munzner P, Hauck CR. Granulocyte CEACAM3 is a phagocytic receptor of the innate immune system that mediates recognition and elimination of human-specific pathogens. *The Journal of experimental medicine* 2004, **199**(1): 35-46.
79. Zebhauser R, Kammerer R, Eisenried A, McLellan A, Moore T, Zimmermann W. Identification of a novel group of evolutionarily conserved members within the rapidly diverging murine Cea family. *Genomics* 2005, **86**(5): 566-580.
80. Kammerer R, Popp T, Singer BB, Schlender J, Zimmermann W. Identification of allelic variants of the bovine immune regulatory molecule CEACAM1 implies a pathogen-driven evolution. *Gene* 2004, **339**: 99-109.
81. Thirunavukarasu P, Sukumar S, Sathaiah M, Mahan M, Pragatheeshwar KD, Pingpank JF, *et al.* C-stage in colon cancer: implications of carcinoembryonic antigen biomarker in staging, prognosis, and management. *Journal of the National Cancer Institute* 2011, **103**(8): 689-697.
82. Chevinsky AH. CEA in tumors of other than colorectal origin. *Seminars in surgical oncology* 1991, **7**(3): 162-166.
83. Chan CH, Stanners CP. Recent advances in the tumour biology of the GPI-anchored carcinoembryonic antigen family members CEACAM5 and CEACAM6. *Current oncology* 2007, **14**(2): 70-73.

84. Thompson JA, Eades-Perner AM, Ditter M, Muller WJ, Zimmermann W. Expression of transgenic carcinoembryonic antigen (CEA) in tumor-prone mice: an animal model for CEA-directed tumor immunotherapy. *International journal of cancer Journal international du cancer* 1997, **72**(1): 197-202.
85. Chan CH, Stanners CP. Novel mouse model for carcinoembryonic antigen-based therapy. *Molecular therapy : the journal of the American Society of Gene Therapy* 2004, **9**(6): 775-785.
86. Chan CH, Cook D, Stanners CP. Increased colon tumor susceptibility in azoxymethane treated CEABAC transgenic mice. *Carcinogenesis* 2006, **27**(9): 1909-1916.
87. Jin L, Li Y, Chen CJ, Sherman MA, Le K, Shively JE. Direct interaction of tumor suppressor CEACAM1 with beta catenin: identification of key residues in the long cytoplasmic domain. *Experimental biology and medicine* 2008, **233**(7): 849-859.
88. Bamberger AM, Riethdorf L, Nollau P, Naumann M, Erdmann I, Gotze J, *et al.* Dysregulated expression of CD66a (BGP, C-CAM), an adhesion molecule of the CEA family, in endometrial cancer. *The American journal of pathology* 1998, **152**(6): 1401-1406.
89. Kunath T, Ordonez-Garcia C, Turbide C, Beauchemin N. Inhibition of colonic tumor cell growth by biliary glycoprotein. *Oncogene* 1995, **11**(11): 2375-2382.
90. Fiori V, Magnani M, Cianfriglia M. The expression and modulation of CEACAM1 and tumor cell transformation. *Annali dell'Istituto superiore di sanita* 2012, **48**(2): 161-171.
91. Gemei M, Mirabelli P, Di Noto R, Corbo C, Iaccarino A, Zamboli A, *et al.* CD66c is a novel marker for colorectal cancer stem cell isolation, and its silencing halts tumor growth in vivo. *Cancer* 2013, **119**(4): 729-738.
92. Lee HS, Ostrowski MA, Gray-Owen SD. CEACAM1 dynamics during neisseria gonorrhoeae suppression of CD4+ T lymphocyte activation. *Journal of immunology* 2008, **180**(10): 6827-6835.
93. Slevogt H, Zabel S, Opitz B, Hocke A, Eitel J, N'Guessan P D, *et al.* CEACAM1 inhibits Toll-like receptor 2-triggered antibacterial responses of human pulmonary epithelial cells. *Nature immunology* 2008, **9**(11): 1270-1278.
94. Hammarstrom S, Baranov V. Is there a role for CEA in innate immunity in the colon? *Trends in microbiology* 2001, **9**(3): 119-125.
95. Rickes S. [Walter Krienitz and one of the first descriptions of gastric bacteria]. *Zeitschrift fur Gastroenterologie* 2011, **49**(11): 1491-1492.
96. Marshall BJ, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1984, **1**(8390): 1311-1315.

97. Marshall BJ. The pathogenesis of non-ulcer dyspepsia. *The Medical journal of Australia* 1985, **143**(7): 319.
98. Goodwin CS, Armstrong JA, Marshall BJ. Campylobacter pyloridis, gastritis, and peptic ulceration. *Journal of clinical pathology* 1986, **39**(4): 353-365.
99. Amieva MR, El-Omar EM. Host-bacterial interactions in *Helicobacter pylori* infection. *Gastroenterology* 2008, **134**(1): 306-323.
100. Oleastro M, Menard A. The Role of *Helicobacter pylori* Outer Membrane Proteins in Adherence and Pathogenesis. *Biology* 2013, **2**(3): 1110-1134.
101. Schistosomes, liver flukes and *Helicobacter pylori*. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. *IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer* 1994, **61**: 1-241.
102. Otero LL, Ruiz VE, Perez Perez GI. *Helicobacter pylori*: The balance between a role as colonizer and pathogen. *Best practice & research Clinical gastroenterology* 2014, **28**(6): 1017-1029.
103. Lertsethtakarn P, Ottemann KM, Hendrixson DR. Motility and chemotaxis in *Campylobacter* and *Helicobacter*. *Annual review of microbiology* 2011, **65**: 389-410.
104. Ottemann KM, Lowenthal AC. *Helicobacter pylori* uses motility for initial colonization and to attain robust infection. *Infect Immun* 2002, **70**(4): 1984-1990.
105. Cao P, Cover TL. Two different families of hopQ alleles in *Helicobacter pylori*. *J Clin Microbiol* 2002, **40**(12): 4504-4511.
106. Boquet P, Ricci V. Intoxication strategy of *Helicobacter pylori* VacA toxin. *Trends in microbiology* 2012, **20**(4): 165-174.
107. Yamaoka Y. Mechanisms of disease: *Helicobacter pylori* virulence factors. *Nature reviews Gastroenterology & hepatology* 2010, **7**(11): 629-641.
108. Odenbreit S, Puls J, Sedlmaier B, Gerland E, Fischer W, Haas R. Translocation of *Helicobacter pylori* CagA into gastric epithelial cells by type IV secretion. *Science* 2000, **287**(5457): 1497-1500.
109. Stein M, Bagnoli F, Halenbeck R, Rappuoli R, Fantl WJ, Covacci A. c-Src/Lyn kinases activate *Helicobacter pylori* CagA through tyrosine phosphorylation of the EPIYA motifs. *Mol Microbiol* 2002, **43**(4): 971-980.
110. Argent RH, Kidd M, Owen RJ, Thomas RJ, Limb MC, Atherton JC. Determinants and consequences of different levels of CagA phosphorylation for clinical isolates of *Helicobacter pylori*. *Gastroenterology* 2004, **127**(2): 514-523.

111. Ren S, Higashi H, Lu H, Azuma T, Hatakeyama M. Structural basis and functional consequence of *Helicobacter pylori* CagA multimerization in cells. *The Journal of biological chemistry* 2006, **281**(43): 32344-32352.
112. Cover TL, Krishna US, Israel DA, Peek RM, Jr. Induction of gastric epithelial cell apoptosis by *Helicobacter pylori* vacuolating cytotoxin. *Cancer research* 2003, **63**(5): 951-957.
113. Gebert B, Fischer W, Weiss E, Hoffmann R, Haas R. *Helicobacter pylori* vacuolating cytotoxin inhibits T lymphocyte activation. *Science* 2003, **301**(5636): 1099-1102.
114. Oldani A, Cormont M, Hofman V, Chiozzi V, Oregioni O, Canonici A, *et al.* *Helicobacter pylori* counteracts the apoptotic action of its VacA toxin by injecting the CagA protein into gastric epithelial cells. *PLoS pathogens* 2009, **5**(10): e1000603.
115. Akada JK, Aoki H, Torigoe Y, Kitagawa T, Kurazono H, Hoshida H, *et al.* *Helicobacter pylori* CagA inhibits endocytosis of cytotoxin VacA in host cells. *Disease models & mechanisms* 2010, **3**(9-10): 605-617.
116. Sheu BS, Yang HB, Yeh YC, Wu JJ. *Helicobacter pylori* colonization of the human gastric epithelium: a bug's first step is a novel target for us. *J Gastroenterol Hepatol* 2010, **25**(1): 26-32.
117. Alm RA, Bina J, Andrews BM, Doig P, Hancock RE, Trust TJ. Comparative genomics of *Helicobacter pylori*: analysis of the outer membrane protein families. *Infect Immun* 2000, **68**(7): 4155-4168.
118. Ilver D, Arnqvist A, Ogren J, Frick IM, Kersulyte D, Incecik ET, *et al.* *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* 1998, **279**(5349): 373-377.
119. Walz A, Odenbreit S, Stuhler K, Wattenberg A, Meyer HE, Mahdavi J, *et al.* Identification of glycoprotein receptors within the human salivary proteome for the lectin-like BabA and SabA adhesins of *Helicobacter pylori* by fluorescence-based 2-D bacterial overlay. *Proteomics* 2009, **9**(6): 1582-1592.
120. Ishijima N, Suzuki M, Ashida H, Ichikawa Y, Kanegae Y, Saito I, *et al.* BabA-mediated adherence is a potentiator of the *Helicobacter pylori* type IV secretion system activity. *The Journal of biological chemistry* 2011, **286**(28): 25256-25264.
121. Voss BJ, Gaddy JA, McDonald WH, Cover TL. Analysis of surface-exposed outer membrane proteins in *Helicobacter pylori*. *J Bacteriol* 2014, **196**(13): 2455-2471.
122. Colbeck JC, Hansen LM, Fong JM, Solnick JV. Genotypic profile of the outer membrane proteins BabA and BabB in clinical isolates of *Helicobacter pylori*. *Infect Immun* 2006, **74**(7): 4375-4378.

123. Mahdavi J, Sonden B, Hurtig M, Olfat FO, Forsberg L, Roche N, *et al.* *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *Science* 2002, **297**(5581): 573-578.
124. Aspholm M, Olfat FO, Norden J, Sonden B, Lundberg C, Sjostrom R, *et al.* SabA is the *H. pylori* hemagglutinin and is polymorphic in binding to sialylated glycans. *PLoS pathogens* 2006, **2**(10): e110.
125. Petersson C, Forsberg M, Aspholm M, Olfat FO, Forslund T, Boren T, *et al.* *Helicobacter pylori* SabA adhesin evokes a strong inflammatory response in human neutrophils which is down-regulated by the neutrophil-activating protein. *Med Microbiol Immunol* 2006, **195**(4): 195-206.
126. Talarico S, Whitefield SE, Fero J, Haas R, Salama NR. Regulation of *Helicobacter pylori* adherence by gene conversion. *Mol Microbiol* 2012.
127. Kao CY, Sheu SM, Sheu BS, Wu JJ. Length of thymidine homopolymeric repeats modulates promoter activity of sabA in *Helicobacter pylori*. *Helicobacter* 2012, **17**(3): 203-209.
128. Rossez Y, Gosset P, Boneca IG, Magalhaes A, Ecobichon C, Reis CA, *et al.* The lacdiNAc-specific adhesin LabA mediates adhesion of *Helicobacter pylori* to human gastric mucosa. *The Journal of infectious diseases* 2014, **210**(8): 1286-1295.
129. Odenbreit S, Faller G, Haas R. Role of the alpAB proteins and lipopolysaccharide in adhesion of *Helicobacter pylori* to human gastric tissue. *International journal of medical microbiology : IJMM* 2002, **292**(3-4): 247-256.
130. Roth A, Mattheis C, Muenzner P, Unemo M, Hauck CR. Innate recognition by neutrophil granulocytes differs between *Neisseria gonorrhoeae* strains causing local or disseminating infections. *Infect Immun* 2013, **81**(7): 2358-2370.
131. Kuespert K, Weibel S, Hauck CR. Profiling of bacterial adhesin--host receptor recognition by soluble immunoglobulin superfamily domains. *J Microbiol Methods* 2007, **68**(3): 478-485.
132. Lee A, O'Rourke J, De Ungria MC, Robertson B, Daskalopoulos G, Dixon MF. A standardized mouse model of *Helicobacter pylori* infection: introducing the Sydney strain. *Gastroenterology* 1997, **112**(4): 1386-1397.
133. Bronsdon MA, Goodwin CS, Sly LI, Chilvers T, Schoenknecht FD. *Helicobacter nemestrinae* sp. nov., a spiral bacterium found in the stomach of a pigtailed macaque (*Macaca nemestrina*). *International journal of systematic bacteriology* 1991, **41**(1): 148-153.
134. JG F, N.S. T, P E, DJ B. *Campylobacter pylori* subsp. *mustelae* subsp. *nov.* isolated from the gastric mucosa of ferrets (*Mustela putorius furo*), and an emended description of *Campylobacter pylori*. *Int J Syst Bacteriol* 1988(38): 367-370.

135. Dailidienė D, Dailidė G, Kersulytė D, Berg DE. Contraselectable streptomycin susceptibility determinant for genetic manipulation and analysis of *Helicobacter pylori*. *Appl Environ Microbiol* 2006 Sep [cited 72 9]5908-5914].
136. Debowski AW, Gauntlett JC, Li H, Liao T, Sehnal M, Nilsson HO, *et al.* Xer-cise in *Helicobacter pylori*: one-step transformation for the construction of markerless gene deletions. *Helicobacter* 2012, **17**(6): 435-443.
137. Alm RA, Trust TJ. Analysis of the genetic diversity of *Helicobacter pylori*: the tale of two genomes. *Journal of molecular medicine* 1999, **77**(12): 834-846.
138. Peck B, Ortkamp M, Diehl KD, Hundt E, Knapp B. Conservation, localization and expression of HopZ, a protein involved in adhesion of *Helicobacter pylori*. *Nucleic acids research* 1999, **27**(16): 3325-3333.
139. Fischer W, Windhager L, Rohrer S, Zeiller M, Karnholz A, Hoffmann R, *et al.* Strain-specific genes of *Helicobacter pylori*: genome evolution driven by a novel type IV secretion system and genomic island transfer. *Nucleic acids research* 2010, **38**(18): 6089-6101.
140. Belogolova E, Bauer B, Pompaiah M, Asakura H, Brinkman V, Ertl C, *et al.* *Helicobacter pylori* outer membrane protein HopQ identified as a novel T4SS-associated virulence factor. *Cellular microbiology* 2013, **15**(11): 1896-1912.
141. Odenbreit S, Kavermann H, Puls J, Haas R. CagA tyrosine phosphorylation and interleukin-8 induction by *Helicobacter pylori* are independent from alpAB, HopZ and bab group outer membrane proteins. *International journal of medical microbiology : IJMM* 2002, **292**(3-4): 257-266.
142. Heuermann D, Haas R. A stable shuttle vector system for efficient genetic complementation of *Helicobacter pylori* strains by transformation and conjugation. *Molecular & general genetics : MGG* 1998, **257**(5): 519-528.
143. Ertl C. Interaktion des Cag-TypIV-Sekretionssystems von *Helicobacter pylori* mit  $\beta$ 1-Integrin zur Untersuchung des Translokationsmechanismus des Cytotoxin-assoziierten Antigens (CagA). *Doctor thesis* 2011.
144. de Jonge R, Pot RG, Loffeld RJ, van Vliet AH, Kuipers EJ, Kusters JG. The functional status of the *Helicobacter pylori* sabB adhesin gene as a putative marker for disease outcome. *Helicobacter* 2004, **9**(2): 158-164.
145. Rieder G, Fischer W, Haas R. Interaction of *Helicobacter pylori* with host cells: function of secreted and translocated molecules. *Curr Opin Microbiol* 2005, **8**(1): 67-73.
146. Fischer W, Puls J, Buhrdorf R, Gebert B, Odenbreit S, Haas R. Systematic mutagenesis of the *Helicobacter pylori* cag pathogenicity island: essential genes

- for CagA translocation in host cells and induction of interleukin-8. *Mol Microbiol* 2001, **42**(5): 1337-1348.
147. Naegele V. Funktionelle Charakterisierung trimerer Autotransporteradhäsine von *Neisseria meningitidis* (NadA) und *Yersinia enterocolitica* (YadA). *Doctor thesis* 2010.
  148. Gray-Owen SD, Dehio C, Haude A, Grunert F, Meyer TF. CD66 carcinoembryonic antigens mediate interactions between Opa-expressing *Neisseria gonorrhoeae* and human polymorphonuclear phagocytes. *EMBO J* 1997, **16**(12): 3435-3445.
  149. Popp A, Dehio C, Grunert F, Meyer TF, Gray-Owen SD. Molecular analysis of *neisserial* Opa protein interactions with the CEA family of receptors: identification of determinants contributing to the differential specificities of binding. *Cellular microbiology* 1999, **1**(2): 169-181.
  150. Hill DJ, Edwards AM, Rowe HA, Virji M. Carcinoembryonic antigen-related cell adhesion molecule (CEACAM)-binding recombinant polypeptide confers protection against infection by respiratory and urogenital pathogens. *Mol Microbiol* 2005, **55**(5): 1515-1527.
  151. Berger CN, Billker O, Meyer TF, Servin AL, Kansau I. Differential recognition of members of the carcinoembryonic antigen family by Afa/Dr adhesins of diffusely adhering *Escherichia coli* (Afa/Dr DAEC). *Mol Microbiol* 2004, **52**(4): 963-983.
  152. Hill DJ, Whittles C, Virji M. A novel group of *Moraxella catarrhalis* UspA proteins mediates cellular adhesion via CEACAMs and vitronectin. *PLoS One* 2012, **7**(9): e45452.
  153. Kupsch EM, Knepper B, Kuroki T, Heuer I, Meyer TF. Variable opacity (Opa) outer membrane proteins account for the cell tropisms displayed by *Neisseria gonorrhoeae* for human leukocytes and epithelial cells. *EMBO J* 1993, **12**(2): 641-650.
  154. Solnick JV, Schauer DB. Emergence of diverse *Helicobacter* species in the pathogenesis of gastric and enterohepatic diseases. *Clinical microbiology reviews* 2001, **14**(1): 59-97.
  155. Boehm MK, Mayans MO, Thornton JD, Begent RH, Keep PA, Perkins SJ. Extended glycoprotein structure of the seven domains in human carcinoembryonic antigen by X-ray and neutron solution scattering and an automated curve fitting procedure: implications for cellular adhesion. *J Mol Biol* 1996, **259**(4): 718-736.
  156. Korotkova N, Yang Y, Le Trong I, Cota E, Demeler B, Marchant J, *et al.* Binding of Dr adhesins of *Escherichia coli* to carcinoembryonic antigen triggers receptor dissociation. *Mol Microbiol* 2008, **67**(2): 420-434.
  157. Voges M, Bachmann V, Kammerer R, Gophna U, Hauck CR. CEACAM1 recognition by bacterial pathogens is species-specific. *BMC Microbiol* 2010, **10**: 117.

158. Farnbacher M, Jahns T, Willrodt D, Daniel R, Haas R, Goesmann A, *et al.* Sequencing, annotation, and comparative genome analysis of the gerbil-adapted *Helicobacter pylori* strain B8. *BMC genomics* 2010, **11**: 335.
159. Beauchemin N, Arabzadeh A. Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. *Cancer metastasis reviews* 2013, **32**(3-4): 643-671.
160. Ahn T, Yim SK, Choi HI, Yun CH. Polyacrylamide gel electrophoresis without a stacking gel: use of amino acids as electrolytes. *Analytical biochemistry* 2001, **291**(2): 300-303.

# Danksagung

---

Mein besonderer Dank gilt Prof. Dr. Rainer Haas für die hervorragende Betreuung dieser Arbeit und die Bereitstellung dieses interessanten Themas. Seine stete Diskussionsbereitschaft und die wertvollen Ratschläge waren maßgeblich am Gelingen dieser Arbeit beteiligt. Vielen Dank für das in mich gesetzte Vertrauen und die Unterstützung. Insbesondere möchte ich mich für die Möglichkeit bedanken, meine Daten in Helsingør vortragen zu dürfen.

Ich möchte auch PD Dr. Wolfgang Fischer für seine Unterstützung und konstruktiven Ideen danken. Sein stets offenes Ohr für Fragen aller Art und die hilfreichen Anregungen haben maßgeblich zum Gelingen dieser Dissertation beigetragen. Vielen Dank auch für das Korrekturlesen dieser Arbeit.

Prof. Dr. Christof Hauck und seinen Mitarbeitern danke ich für die engagierte Kooperation und die mir zur Verfügung gestellten Tools. Prof. Dr. Dr. Heesemann danke ich für die Bereitstellung von CEACAM-transfizierten HEK293 Zellen.

Ich möchte mich bei allen „Haasen“ und Nicht-„Haasen“ für die angenehme Arbeitsatmosphäre und die Diskussions- und Hilfsbereitschaft bedanken. Besonders herzlich erinnere ich mich an die unzähligen, ausgelassenen Momente in der Mittagsrunde und während der Kaffeepausen. Vielen Dank, Friederike, für Rat und Tat in allen Lebensangelegenheiten!

Liebe Freunde, ich möchte mich bei euch für eure optimistische Unterstützung und die unzähligen, schönen Stunden während der vergangenen Jahre bedanken. Ihr habt mein Leben sehr bereichert und mir geholfen, das Wesentliche nicht aus den Augen zu verlieren. In Zukunft werde ich wieder mehr Zeit für Euch haben!

Liebe Laura, danke für deine bedingungslose Freundschaft und die vielen tollen Dinge, die wir in den letzten Jahren zusammen unternommen haben. Du bist zu meiner besten Freundin geworden, meiner dritten Schwester. Ich bin froh, dass es dich gibt.

Liebes Lab66, auch wenn es gefühlt schon ewig her ist, ihr habt mir soviel beigebracht. Was wäre ich ohne euch!

Mein besonderer Dank gilt meiner Familie, die mir in jeder Lebenslage Halt gegeben hat. Danke für eure Liebe und bedingungslose Unterstützung in jeglicher Hinsicht. Ich weiß, ich kann immer auf euch zählen.

