

Quantitative analysis of Neisserial Opa protein interactions with human CEACAMs *in vitro*.

Jennifer Nicole Martin
Lancaster, PA

B.S. The Pennsylvania State University, Berks College, 2011

A dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Chemistry

University of Virginia
May, 2016

© Copyright by
Jennifer Nicole Martin
All rights reserved.
May 2016

Abstract

Increasing bacterial resistance to antibiotics is an urgent global health concern. In order to combat this growing antibiotic resistance, alternative treatments must be developed. In order to develop new treatments, a deeper understanding of bacterial pathogenesis is needed. This dissertation presents investigations of the interaction between Opa proteins from *Neisseria gonorrhoeae* and *Neisseria meningitidis*, pathogens which cause gonorrhea and meningococcal meningitis respectively, and the human receptor CEACAM.

Opa proteins are a family of 8-stranded β -barrel proteins found in the outer membrane of pathogenic *Neisseria*. Opa proteins contain regions of high sequence variation, which are involved in receptor interactions. The majority of Opa proteins interact with carcino-embryonic antigen-like cell adhesion molecules (CEACAMs), expressed on human host cell surfaces. Interactions of Opa proteins with CEACAM receptors are able to trigger engulfment of the bacterium into the host cell.

While Opa proteins appear to selectively engage specific receptors, the molecular determinants of the interaction are unknown. Multiple sequence alignments of Opa protein sequences have not revealed a conserved receptor recognition motif. The investigations presented in this dissertation seek to gain a deeper understanding of Opa protein – CEACAM receptor interactions. Towards this end, an *in vitro* system was developed containing recombinant Opa proteins folded into liposomes.

A selection of CEACAM binding Opa proteins were investigated for their ability to interact with the recombinant N-terminal domain of CEACAM. Binding was qualitatively assessed using pull-down assays and immunoblotting to compare the specificity of Opa proteins *in vitro* to their *in vivo* function, and validate the Opa proteoliposome system for use

in further experiments. Quantitative experiments were conducted to investigate the kinetics and thermodynamics of the Opa protein – CEACAM receptor interaction. A wider selection of Opa proteins were investigated for their affinity towards various CEACAM receptors. Tight affinities were calculated for all investigated interactions.

Knowledge of Opa – CEACAM affinities is the first step towards gaining a deeper understanding of how *Neisseria* are able to trigger engulfment into host cells. Phagocytosis of the Opa proteoliposomes is being investigated to determine if there is a correlation between receptor specificity and bacterial engulfment. Various techniques are being utilized to gain insight into the mechanism of molecular recognition between Opa proteins and CEACAM receptors. Elucidating the molecular determinants of the Opa protein – CEACAM receptor interactions will provide the foundation for the design of a targeted liposome therapeutic delivery system for human cells.

Acknowledgements

I am forever thankful for the support of Dr. Linda Columbus. I could not have asked for a better advisor for my graduate career. Thank you for seeing potential in me, for all of the opportunities to grow, for motivating me to always better myself, and for encouraging and supporting my interest in both science and academia.

Thank you to all of the Columbus lab members, both past and present for their support, scientific discussions, and friendship. In particular, Dr. Alison Dewald, Dr. Dan Fox and Dr. Ryan Lo for training me, helpful discussions, support, and friendship which has continued after their departure from the lab. Thank you to Ashton Brock, Marissa Kieber, Jason Kuhn, Steven Keller, Nicole Swope, and Tracy Caldwell. Thank you to Cameron Mura and the members of the Mura lab for their helpful discussions, technical assistance, and friendship. Thank you to Peter Randolph and Kimberly Stanek in particular.

Thank you to our collaborators, Dr. Alison Criss and Louise Ball, without whom this work would not be possible. I am lucky to have been able to work with such collaborators who are not only great scientists, but wonderful people as well.

Thank you to Dr. Carol Price, for helpful scientific discussion, and for her friendship and support which has meant more to me than she knows. I am forever grateful to her for always being there for me, in times of both joy and frustration.

And finally, thank you to my friends and family, for their unwavering support of me in everything that I do.

Table of Contents

Copyright page	ii
Abstract.....	iii
Acknowledgements	v
Table of contents	vi
List of figures.....	xi
List of Tables	xiv
 Chapter 1: Introduction	 1
1.1 Emerging antibiotic resistance of bacterial pathogens is an urgent global health problem	1
1.2 Neisserial opacity associated (Opa) proteins	4
1.2.1 Background and significance	4
1.2.2 Opa protein structure and function	5
1.3 Carcino-embryonic antigen-like cell adhesion molecules	12
1.3.1 Biological importance	12
1.3.2 CEACAMs act as a receptor for Opa proteins	17
1.3.3 CEACAM mediated engulfment of Opa expressing bacteria	19
1.4 Liposomes are a useful tool for studying membrane proteins <i>in vitro</i>	20
1.5 Dissertation overview	21
1.6 References	23

Chapter 2: Quantitative methods used to investigate the Opa–CEACAM interaction	..33
2.1 Overview33
2.2 Selection of lipid composition for <i>in vitro</i> investigations of Opa-CEACAM interactions33
2.2.1 Membrane proteins in research33
2.2.2 Utilizing liposomes to investigate membrane proteins36
2.2.3 Developing a liposomal system for Opa proteins37
2.2.4 Selection of PEGylated lipids39
2.2.5 Trypsin cleavage of extracellular Opa protein loops42
2.3 Understanding the thermodynamics of protein – ligand interactions45
2.4 Fluorescence polarization (FP) is well suited to study Opa – CEACAM interactions46
2.4.1 Underlying principles of FP46
2.4.2 Opa – CEACAM interactions can be assayed using FP50
2.4.3 Analysis of FP data can be complicated by several factors52
2.5 Biolayer interferometry (BLI) provides fast and label free assessment of interactions53
2.5.1 Underlying principles of BLI53
2.5.2 Assessing Opa – CEACAM interactions using BLI58
2.6 References61

Chapter 3: Reconstituted Opa proteoliposomes are an effective tool for studying Opa protein interactions <i>in vitro</i>	67
3.1 Overview	67
3.2 Introduction	68
3.2.1 Neisserial attachment to host cells	68
3.2.2 CEACAM as an adhesion molecule in cells	68
3.2.3 CEACAMs act as a receptor for Neisserial Opa proteins	69
3.3 Results and discussion	75
3.3.1 Recombinant NCCM1 and 3 interact with Opa ₆₀ and OpaD+ Gc	75
3.3.2 Recombinant Opa ₆₀ and OpaD reconstituted in liposomes interact with recombinant NCCM1 and 3	78
3.3.3 NCCM1 and 3 interact with Opa ₆₀ and OpaD with nanomolar affinity	80
3.3.4 Competition for CEACAM may explain high affinities	85
3.4 Concluding remarks	86
3.5 References	87
 Chapter 4: Binding affinities of a variety of Opa proteins for CEACAMs <i>in vitro</i> suggests the interaction may be non-specific	93
4.1 Overview	93
4.2 Introduction	94
4.2.1 Opa _{CEA} selectivity <i>in vivo</i>	94

4.2.2 Mechanisms of CEACAM induced bacterial internalization	98
4.3 Results and discussion	103
4.3.1 The Opa – CEACAM interaction is of high affinity for all combinations investigated.....	103
4.3.2 Varying the amount or molecular weight of PEGylated lipids has little to no effect on Opa – CEACAM interactions.....	111
4.3.3 The lipid:Opa ratio has little effect on Opa association with NCCM1	114
4.3.4 Understanding Opa – CEACAM selectivity	118
4.3.5 Opa – CEACAM interactions appear to be monovalent	122
4.4 Concluding remarks.....	123
4.5 References	124
 Chapter 5: Prospects for future research into the molecular determinants of Opa – CEACAM interactions	 133
5.1 Overview	133
5.2 Mapping the regions of Opa HV1 and HV2 that interact with CEACAM.....	133
5.2.1 Competition assays using FP to determine which regions on Opa HV loops are involved in receptor interactions	134
5.2.2 Protein footprinting to assess which amino acid residues on the loops of Opa proteins are involved in binding to receptors	139

5.3 Determining the structure of the Opa – CEACAM complex	144
5.3.1 Progress towards structure determination	144
5.3.2 Prospects for future research	145
5.4 Long term applications of this research	148
5.5 Concluding remarks.....	148
5.6 References	150
 Appendix: Materials and methods.....	 153

List of Figures

Chapter 1

1.1 Micrograph of Gc colonies	8
1.2 Micrograph of <i>Neisseria gonorrhoeae</i> invading human cells	9
1.3 NMR structure of Opa ₆₀ , an Opa _{CEA} from Gc.....	10
1.4 Sequence alignment of the N-terminal domain of CEACAMs	15
1.5 Domain organization of CEACAMs.....	16
1.6 Cartoon representation of the N-terminal domain of human CEACAM1 and homology model of CEACAM3	18

Chapter 2

2.1 Progress in membrane protein structure determination	35
2.2 Possible conformations of PEG attached to a membrane structure	41
2.3 Trypsin cleavage of Opa proteoliposomes.....	44
2.4 The Jablonski energy diagram explains the principles of fluorescence.....	48
2.5 The principles of fluorescence polarization.....	49
2.6 Schematic of FP with CEACAM and Opa proteoliposomes	51
2.7 Interference patterns of waves	55
2.8 Representative surface of a biosensor.....	56
2.9 Schematic of biolayer interferometry	57
2.10 Theoretical data from a BLI experiment with GST-NCCM and Opa proteoliposomes ..	60

Chapter 3

3.1 Surface representation of NCCM1.....	72
3.2 Sequence alignment of human NCCMs 1, 3, 4, 5, 6, 7, and 8.....	73
3.3 Sequence alignment of Opa ₆₀ from Gc strain MS11, OpaD from Gc strain FA1090, and Opa ₅₀ from Gc strain MS11	74
3.4 CEACAMs interact specifically with Opa _{CEA} proteins expressed in Gc.....	76
3.5 Immunoblot of pull-down assay using Opa expressing Gc and GST-NCCM1	77
3.6 Recombinant Opa _{CEA} proteins retain their ability to interact with CEACAM	79
3.7 MALDI spectra of fluorescently labeled (fl-) NCCM1 and NCCM3.....	81
3.8 Structure of NCCM1 H139C	82
3.9 Opa _{CEA} proteins in liposomes interact with NCCM with a nanomolar affinity.....	83
3.10 Competition FP experiments with unlabeled NCCMs 1 and 3.....	84

Chapter 4

4.1 <i>Neisseria gonorrhoeae</i> N313 (Opa ₅₇) triggers different structural modifications on HeLa cell surfaces, depending on the expression of CEACAM1 or CEACAM3	102
4.2 Representative raw BLI binding data	108
4.3 CEACAM1+ HeLa cells show higher internalization of Opa ₆₀ proteoliposomes than CEACAM- cells.....	121

Chapter 5

5.1 Sequences of the Opa ₆₀ HV regions used for competition experiments	137
5.2 Peptide competition experiments indicate decreased Opa – CEACAM interactions with peptide present	138
5.3 Preliminary protease protection results using trypsin	141
5.4 Cleavage patterns of Opa ₆₀ loops.....	142
5.5 Theoretical mass spectrometry data for trypsin cleavage of Opa proteoliposomes with and without NCCM.....	143
5.6 A selection of preliminary crystallization leads.....	146
5.7 Crystal packing of NCCM1	147

List of tables

Chapter 1

1.1 Examples of attachment sites and maladies of invasive bacteria	3
1.2 Comparison of Gc MS11 Opa protein nomenclature	7
1.3 Sequence alignment of the HV regions of Opa proteins.....	11

Chapter 4

4.1 Receptor specificity of Opa proteins from Gc strain MS11	96
4.2 HV sequences of Opa proteins from Gc strain MS11.....	97
4.3 Kinetic and thermodynamic parameters for Opa _{CEA} interactions with a variety of NCCM receptors	109
4.4 Kinetic and thermodynamic parameters for non-Opa _{CEA} proteins interacting with a variety of NCCM receptors.....	110
4.5 Kinetic and thermodynamic parameters for the Opa ₆₀ – NCCM1 interaction with varied amounts of PEG1000 or PEG2000 in liposomes	113
4.6 Kinetic and thermodynamic parameters for the Opa ₆₀ – NCCM1 interaction with varied lipid:Opa ratios.....	116
4.7 Liposome size and polydispersity with various lipid:Opa ratios	117

1. Introduction

1.1 Emerging antibiotic resistance of bacterial pathogens is an urgent global health problem

Bacterial pathogens have long been a major cause of global health issues. These pathogens are extremely diverse and are the causative agents of a variety of diseases, ranging from multiple different sexually transmitted infections (STIs) to the plague (Table 1.1). Traditionally, the primary method of treatment for many of these bacterial infections is the use of antibiotics. However, an increasing number of bacteria are becoming resistant to many common antibiotics, and some bacterial strains that have become resistant to all known antibiotics to date.¹⁻³

In the United States alone, there are over 2 million people annually with new infections from bacteria that are resistant to at least one antibiotic, and at least 23,000 deaths are a direct result of these infections.¹ Antibiotic resistance is a global threat, as new forms of antibiotic resistance can easily spread between continents. Among all of the bacterial resistant problems, gram-negative bacteria are most concerning, as they are becoming resistant to far more antibiotics than gram-positive bacteria.¹ In 2013 the United States Center for Disease Control and Prevention (CDC) released a report detailing the complex problem of antibiotic resistance, summarizing the threats from numerous drug resistant bacteria, such as *Clostridium difficile*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, and *Neisseria gonorrhoeae*.¹

Of particular interest to the Columbus lab and this dissertation are the pathogenic Neisserial species. *Neisseria gonorrhoeae* (also called the gonococcus, or Gc) and *Neisseria meningitidis* (also called the meningococcus, or Nm), the causative agents of gonorrhea and meningococcal meningitis, respectively, are gram-negative human specific pathogens that are

able to colonize and invade mucosal tissue. Meningococcal meningitis infects the mucosal surface of the nasopharynx, is the most common form of bacterial meningitis, and is the causative agent of septicemia.⁴ From 2013 – 2015 there were four outbreaks of meningitis on college campuses in the United States, which resulted in the death of two students as well as multiple reports of students experiencing negative neurological effects.⁵

Gonorrhea is an infection of the mucosal surface of the genitor-urinary tract, and is the second most common STI in the United States.¹ Gc is becoming increasingly resistant to antibiotics, and in 2013 the CDC classified antibiotic resistant Gc as an urgent threat to global public health.¹ The CDC estimates that there will be 820,000 new Gc infections in the US every year, and approximately 30% of those cases will be resistant to any antibiotic.¹ Even more concerning than Gc which are resistant to a specific antibiotic, there has been a “superbug” strain of Gc identified that is resistant to all common antibiotics available today.²

The emergence of these antibiotic strains of invasive pathogenic bacteria calls for new methods of treatment for these diseases. In order to develop new treatment techniques, a deeper understanding of bacterial colonization and infection is essential.

Table 1.1: Examples of attachment sites and maladies of invasive bacteria. Adapted from Todar's Online Textbook of Bacteriology.⁶

Bacterium	Attachment site	Disease
<i>Streptococcus pyogenes</i> ^c	Pharyngeal epithelium	Sore throat
<i>Streptococcus mutans</i> ^c	Pellicle of tooth	Tooth decay
<i>Streptococcus pneumoniae</i> ^{b, c}	Mucosal epithelium	Pneumonia
<i>Staphylococcus aureus</i> ^b	Mucosal epithelium	Various
<i>Neisseria gonorrhoeae</i> ^a	Urethral / cervical epithelium	Gonorrhea
<i>Enterotoxigenic Escherichia coli</i> ^a	Intestinal epithelium	Diarrhea
Uropathogenic <i>E. coli</i> ^a	Urethral epithelium or Upper urinary tract	Urethritis or Pyelonephritis
<i>Bordetella pertussis</i>	Respiratory epithelium	Whooping cough
<i>Vibrio cholerae</i>	Intestinal epithelium	Cholera
<i>Treponema pallidum</i>	Mucosal epithelium	Syphilis
<i>Mycoplasma</i>	Respiratory epithelium	Pneumonia
<i>Chlamydia</i>	Conjunctival or urethral epithelium	Conjunctivitis or urethritis

^a Indicates antibacterial resistant strains are an urgent threat¹

^b Indicates antibacterial resistant strains are a serious threat¹

^c Indicates antibacterial resistant strains are a concerning threat¹

1.2 Neisserial opacity associated (Opa) proteins

The mechanism of colonization or invasion of a bacterial pathogen into a host cell often involve interactions between membrane proteins on surfaces of host cells and the bacteria. In the case of the pathogenic *Neisseria* species, this interaction can be facilitated by opacity associated (Opa) proteins. Following initial *Neisseria* attachment, which is mediated by pili and adhesins, Opa proteins interact with human cell receptors, and can trigger the engulfment of the bacterium by host cells, even by host cells that do not normally undergo phagocytosis (Fig. 1.1).⁷⁻⁹

1.2.1 Background and significance

Opa proteins are a family of outer membrane proteins found in the pathogenic *N. gonorrhoeae* (Gc) and *N. meningitidis* (Nm). First identified in Gc, Opa proteins are named as such because their expression changes the color and opacity of the bacterial colonies (Figure 1.2).¹⁰⁻¹² This opaque phenotype is thought to arise because of Opa-induced bacterial adhesion and aggregation.¹³ Meunsner and colleagues demonstrated that laboratory strains of *E. coli* cells transformed to express Opa proteins to the outer membrane were engulfed by HeLa cells transfected to express the appropriate receptor.¹⁴ This suggests that Opa proteins alone are sufficient to trigger bacterial engulfment.

Opa proteins are encoded by a family of unlinked genes; there are eleven genes in Gc, and up to four in Nm (Table 1.2).¹¹ A high level of diversity exists in Opa protein expression and sequence. Even though there are approximately 15 different *opa loci*, many more Opa protein sequences have been observed. This sequence diversity stems primarily from recombination, with some contribution from point mutations as well.^{9,15} *Neisseria* are highly

competent bacteria, which enables them to uptake exogenous DNA and incorporate it into their own genome.⁹ This allows for full and partial *opa* genes to recombine between loci of a single organism, as well as between loci of other organisms as well.

In addition to such high sequence diversity, Opa protein expression can be turned on or off through translational phase variation as well. All *opa* genes contain an amino-terminal leader peptide containing tandem repeats of the DNA sequence [CTCTT]_n.¹⁶ The number of pentameric repeats in the mRNA is necessary for correct translation of the mature Opa protein. Due to slipped-strand mispairing during translation, the number of repeats can change. This variation can shift the sequence out of the reading frame, causing expression of Opa proteins to be turned off.¹⁶ Because of the phase variable expression, *Neisseria* may express zero, one, or multiple different Opa variants in a single bacterium at any given time.

1.2.2 *Opa protein structure and function*

The Opa protein structure consists of a membrane spanning eight stranded β -barrel, linked by four extracellular loops (Fig. 1.3).¹⁷ Within the Opa family, the barrel of the protein and the short extracellular loop 4 have a highly conserved sequence (approximately 70% sequence identity).¹⁸ Extracellular loop 1 has a region that exhibits slight sequence diversity (the semi-variable loop, or SV, Fig. 1.3 in yellow), while the extracellular loops 2 and 3 have regions of high sequence diversity (the hypervariable loops, HV1 and HV2, respectively, Fig. 1.3 in red, Table 1.3).⁹ Receptor specificity is determined primarily by the HV1 and HV2 sequences.¹⁹ However, the specific amino acids in these HV regions responsible for interacting with receptors has yet to be determined. The sequences of the Opa HV regions are so diverse that multiple sequence alignment does not reveal consensus motifs. In addition, Bos and

colleagues investigated Opa - CEACAM interactions using chimeric Opa proteins which contained different HV1 and HV2 loop sequences from a variety of Opa_{CEA} proteins. Receptor specificity was abolished in the chimeric Opa proteins, which demonstrates that the Opa protein primary sequence is insufficient for receptor recognition and engagement; it is specific combinations of HV1 and HV2 sequences that are necessary for interaction with the appropriate receptor (described in detail in chapter 4).²⁰

To date, there are 26 SV sequences, 97 HV1 sequences, and 127 HV2 sequences found from the 345 unique *opa* alleles sequenced. Interestingly, different combinations of these HV and SV sequences could yield over 300,000 possible Opa sequences, yet only a minor fraction of that has been observed. This observation may suggest that some combinations of sequences results in either unstable or functionally inactive proteins.

Opa proteins are known to engage two primary receptors, depending on the HV sequences, and can be classified according to their binding partner. Opa_{HS} bind to heparan sulfate proteoglycans (HSPGs) or integrin receptors via a heparan-mediated interaction with fibronectin or vitronectin. The more abundant class, Opa_{CEA}, bind members of the carcino-embryonic antigen-like cellular adhesion molecule (CEACAM, or CCM) family. Typically, initial *Neisseria* attachment is mediated by pili and adhesion proteins.⁸ Following this initial attachment, Opa proteins are interact with receptors; however, the interaction of Opa proteins with receptors alone is sufficient to drive adhesion and engulfment into host cells.¹⁴ Additional information on Opa protein receptor specificity and binding is found in Chapter 4.

Table 1.2 Comparison of Gc MS11 Opa protein nomenclatures. Adapted from Bos *et al.*¹³

Opa protein^a	<i>opa</i> locus^b	rOpa^c
OpaA	<i>C</i>	Opa ₅₀
OpaB	<i>K</i>	Opa ₅₇
OpaC	<i>G</i>	Opa ₅₂
OpaD	<i>F</i>	Opa ₅₆
OpaE	<i>E</i>	Opa ₅₅
OpaF	<i>I</i>	Opa ₅₄
OpaG	<i>J</i>	Opa ₅₈
OpaH	<i>D</i>	Opa ₅₉
OpaI	<i>H</i>	Opa ₆₀
OpaJ	<i>B</i>	Opa ₅₁
OpaK	<i>A</i>	Opa ₅₃

^a Designation of naturally occurring Opa variants from Swanson *et al.*²¹

^b Designation of chromosomal *opa* loci from Bhat *et al.*²²

^c Designation of recombinant Opa variants from Kupsch *et al.*²³

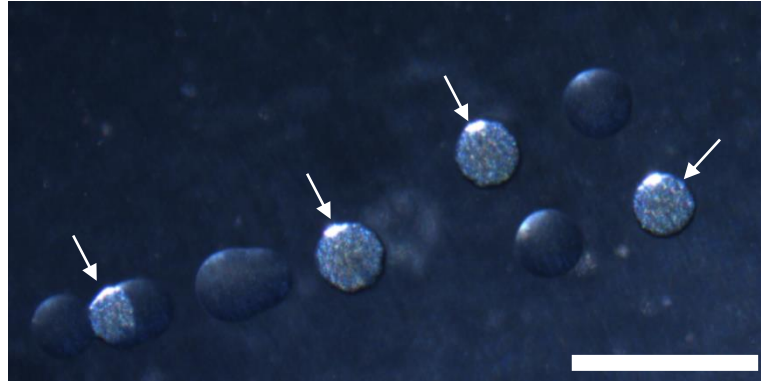


Figure 1.1: Micrograph of Gc colonies. *Photo courtesy of Louise M. Ball.* A photograph taken under a light microscope with a mirrored substage depicting Gc colonies either expressing Opa proteins (opaque colonies indicated with white arrows), or not expressing Opa proteins (smooth, transparent colonies). Scale bar = 1 mm.

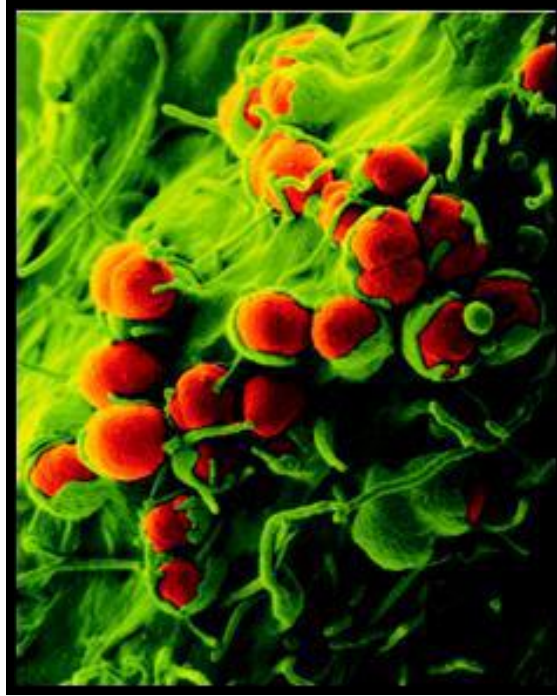


Figure 1.2: Micrograph of *Neisseria gonorrhoeae* invading human cells. *Cover image of Bilker et al, reprinted with permission.*²⁴ A falsely colored scanning electron micrograph depicting Gc (red) being engulfed by HeLa cells (green), which have been transfected to express CEACAM1, which is a receptor for Op_{CEA} proteins.

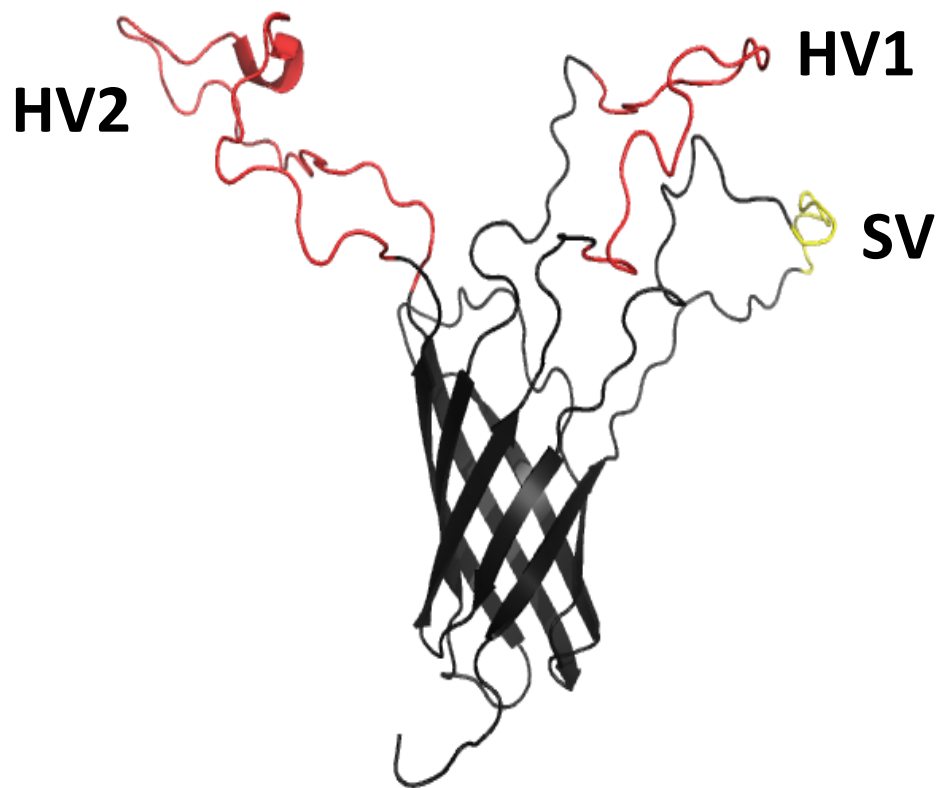


Figure 1.3: NMR structure of Opa₆₀, an Opa_{CEA} from Gc. (PDB IDs 2MLH, 2MAF¹⁷) The structure of Opa proteins consists of an 8-stranded β -barrel (black) with four extracellular loops. The loops are classified based on sequence homology among members of the Opa family. Loop one contains a semi-variable region (SV, yellow), loops two and three contain hypervariable regions (HV1 and HV2, red), and loop 4 is conserved (black).

Table 1.3: Sequence alignment of the HV regions of a selection of Opa proteins.

HV1 Sequence	HV2 Sequence	CEACAM				HSPG
		1	3	5	6	
NDNKYSVNTKNVQVNKSNGN-RQD-----LK	KHQVHSVESKTTIVTSKPTKGATQPGKLVSGPTPKPAYHESNSISSL	√	√	√	√	X
NDNKY.VNT+N*QVNKSNGN-RQD-----*+	+ .S*D.*K-K..N*LTVPNIPGGTPT*YNQGSTQDA...H.*RRL	√	X	√	√	X
NDNKY.VNT+N*QVNKSNGN-RQD-----*+	+ .Q*H.*KKE..T*FLAPTGDAKVP GK*VEGPFSKPA...H.*SSL	√	X	√	√	X
NNNYY.VNT+E*RRNNNAGN-WRE-----*+	+ .Q*H.*KKE..T*FLAPTGDAKVP GK*VEGPFSKPA...H.*SSL	√	X	√	X	X
NNNKY.VNT+E*QRNNSNGTTWKE-----*+	+ .Q*H.*RKE..T*FSKPSGSTTKPGE*PS-LVTKPA...N.*SSL	√	X	√	X	X
NNNKY.VNT+E*ENKNNNK---RD-----*+	+ .S*D.*KKT.EV*TSSHA--PGTAPT*YNVPKTQNA...H.*RRL	√	X	√	X	X
KESNY.---+K*TEFKHQNGNKQE-----D+	+ .Q*H.*ESK..I*TSKPTKGATQPGK*VSGPTPKPA...N.*SSL	√	X	√	X	X
KESNY.---+K*TEFKHQNGNKQE-----D+	+ .Q*H.*ESK..T*TTN--NGGEVP----QGPTPKPA...H.*SSV	√	X	X	X	X
KESNS.T--+K*TEEINNN--YKE-----TQ	+ .Q*H.*ETK..T*TSKPKNGSPQGGP*IQTDPSKPP...H.*SSV	√	X	X	X	X
KESNF.T--+K*TEEIKDN--YKE-----T+	+ .Q*H.*ETK..T*TSKPKGGTPAGGP*IKTDPSKPP...H.*SSL	√	X	X	X	X
NNNKY.VNT+N*QKND-NGN-RQD-----*+	+ .S*D.*KKT..N*VTVAGAANTAPT-*YYAPETQNA...H.*RRL	X	X	X	X	ND
SDNKY.VSI+NMRVHKHNSN-RKN-----*+	+ .S*D.*KKI.GL*TTST-PGIMSG--*YKVLRTPGAHR..D.*RRV	X	X	X	X	√
NNSKY.VSI+E*GRNDNSTS-NSSHLNIKTQ+	+ .Q*R.*ESE..T*TTH--NGAP----*PQGPTPKPA..K.R.*SSL	X	X	X	X	√

+ conservation of charge

* conservation of hydrophobic residue

. conservation of polar residue

- indicates a gap

ND not determined

1.3 Carcino-embryonic antigen-like cell adhesion molecules

1.3.1 Biological importance

CEACAMs (Carcino-Embryonic Antigen-like Cellular Adhesion Molecules) are a subgroup of the immunoglobulin superfamily involved in many cellular processes such as cell adhesion, proliferation, differentiation, and tumor suppression. Related to these various functions, CEACAM dysregulation is often observed in implantation of circulating tumors²⁵ and tumor angiogenesis.²⁶ To date, there have been twelve different CEACAM variants identified in humans: CEACAM1, CEACAMs 3-8, CEACAM16, and CEACAMs 18-21.^{27,28} Of these CEACAMs, all are exclusively expressed in humans except for CEACAM16 and CEACAMs 18-20, which are also expressed in mice.²⁷ CEACAM1, and CEACAMs 3-8 interact with pathogens, whereas CEACAM16, and CEACAMs 18-21, which were discovered much more recently, have not yet been investigated.²⁷ As such, this dissertation will focus on CEACAM1, and CEACAMs 3-8.

CEACAMs have long been of interest because of their involvement in tumors, but studies in the last decade focus more on their contribution in other physiological processes. The carcinoembryonic antigen (CEA) was first identified in the mid-1960s, later renamed to CEACAM5.²⁷ CEACAM5 was discovered as an important tumor-associated antigen in colon cancer, and was later found to be involved in cell-cell adhesion as well.^{27,29} Overexpression of CEACAM5, and of other CEACAMs as well, is associated with abnormal growth of adherent cells.³⁰

CEACAM1 is the most investigated member of the CEACAM family, likely because it is highly expressed in rodents, making *in vivo* experimentation more accessible.²⁷ CEACAM1 contributes to cell-cell adhesions via homo- and heterotypic interactions.²⁷

Additionally, CEACAM1 is important in signal transduction, where it is involved in angiogenesis as an effector of the vascular endothelial growth factor (VEGF), insulin regulation via signaling pathways initiated by the insulin receptor, tumorigenesis as a suppressor of cell proliferation and a promoter of apoptosis, and bacterial pathogenesis as a receptor that promotes adhesion and engulfment.³¹⁻³⁴ In contrast to CEACAM1, which has been implicated in numerous biological functions, CEACAM3 is thought to have evolved as innate immune protection. CEACAM3 has no known endogenous ligand; it interacts specifically with proteins expressed on the surface of human-specific bacterial pathogens, and these interactions mediate uptake of the pathogen that often induces killing of the bacterium via oxidative burst and toxic granule release.^{27,35}

CEACAM expression varies among members of the protein family. CEACAM1 has the widest tissue distribution among the characterized family members, including constitutive expression in epithelium and leukocytes as well as inducible expression in endothelial cells, T-cells, neutrophils, and lymphocytes.²⁷ In contrast, CEACAM3 is expressed exclusively on granulocytes.^{36,37} CEACAM4 is not as well understood as other members of the CEACAM family, most likely because to date, an endogenous ligand that interacts with CEACAM4 has not been identified. However, based on cDNA library analysis, CEACAM4 is only expressed in myeloid cells.³⁸ Normal expression of CEACAM5 is found only in epithelial cells, most abundantly on the apical surface of the gastrointestinal tract, but it is present on other mucosal epithelia as well (nasopharynx, lung, urogenital tract, and sweat glands).³⁹ CEACAM6 has broad tissue distribution, including the epithelia of numerous body organs, as well as granulocytes.⁴⁰ Expression of CEACAM7 is limited to the apical surface of the colon epithelium.⁴⁰ Similar to CEACAM3, CEACAM8 is only expressed by granulocytes.⁴¹ Since

CEACAMs are so broadly distributed throughout human tissue, it can allow pathogenic *Neisseria* to use interactions with CEACAMs for initial invasion and transcellular travel to sub-epithelial locations.⁴²

The extracellular regions of all CEACAMs consist of one highly conserved amino-terminal immunoglobulin variable (IgV)-like domain (Fig. 1.4, sequences; Fig. 1.5, blue), and one to six immunoglobulin constant (IgC)-like domains (Fig. 1.5, green).^{9,10} There are two types of IgC domains: type A, which consists of 93 amino acids, and type B, which consists of 85 amino acids. CEACAMs can be embedded in the membrane via two different structural elements: (1) a transmembrane hydrophobic domain that spans the lipid bilayer, and connects to a cytoplasmic domain, or (2) a glycosylphosphatidylinositol (GPI) element that anchors the protein to the lipid bilayer, and lacks a cytoplasmic domain. CEACAMs 1, 3, and 4 contain a transmembrane domain (Fig. 1.5, purple) while CEACAMs 5-8 contain a GPI anchor (Fig. 1.5, red). The transmembrane domains of CEACAM1 and 3 are involved in signal transduction, while the GPI anchors of CEACAMs 5-8 are known to direct membrane proteins to cholesterol and sphingolipid rich micro-domains in the cell membrane.^{27,43}

```

          40          50          60          70          80          90
CEACAM1 QLTTESMPFNVAEGKEVLLLVHNLPPQLFGYSWYKGERVDGNRQIVGYAIGT-QQATPGP
CEACAM3 KLTIESM.LSV....E....VH.LPQHLLFG.S.Y..ERVDGNSL.V..VIGT-.QAT..A
CEACAM4 QFTIEAL.SSA....D....AC.ISETIQA.Y.H..KTAEGSPL.A..ITDI-.ANI..A
CEACAM5 KLTIEST.FNV....E....VH.LPQHLLFG.S.Y..ERVDGNRQ.I..VIGT-.QAT..P
CEACAM6 KLTIEST.FNV....E....AH.LPQNRIG.S.Y..ERVDGNSL.V..VIGT-.QAT..P
CEACAM7 QTNIDVV.FNV....E....VH.ESQONLYG.N.Y..ERVHANYR.I..VKNIS.ENA..P
CEACAM8 QLTIEAV.SNA....E....VH.LPQDPRG.N.Y..ETVDANRR.I..VISN-.QIT..P

          100          110          120          130          140
CEACAM1 ANSGRETIYPNASLLIQNVNTQNDTGIFYTLQVIKSDLVNNEEATGQFHVYP
CEACAM3 .YSG..TI.T.AS..IQ.VTQN.I.F...QV.KSDLVNNEEA.G.FH.YQ
CEACAM4 .YSG..TV.P.GS..FQ.ITLE.A.S...RT.NASYDSQA.G.LH.HQ
CEACAM5 .YSG..II.P.AS..IQ.IIQN.T.F...HV.KSDLVNNEEA.G.FR.YP
CEACAM6 .YSG..TI.P.AS..IQ.VTQN.T.F...QV.KSDLVNNEEA.G.FH.YP
CEACAM7 .HNG..TI.P.GT..IQ.VTHN.A.F...HV.KENLVNEEV.R.FY.FS
CEACAM8 .YSN..TI.P.AS..MR.VTRN.T.S...QV.KLNLMSSEV.G.FS.HP

```

Figure 1.4: Sequence alignment of the N-terminal domains of CEACAMs. Numbering is based on the full-length CEACAM1 sequence (UniProt ID P13688-1). Opa-binding CEACAMs are labeled in bold font. . represents a conserved residues, - represents a gap.

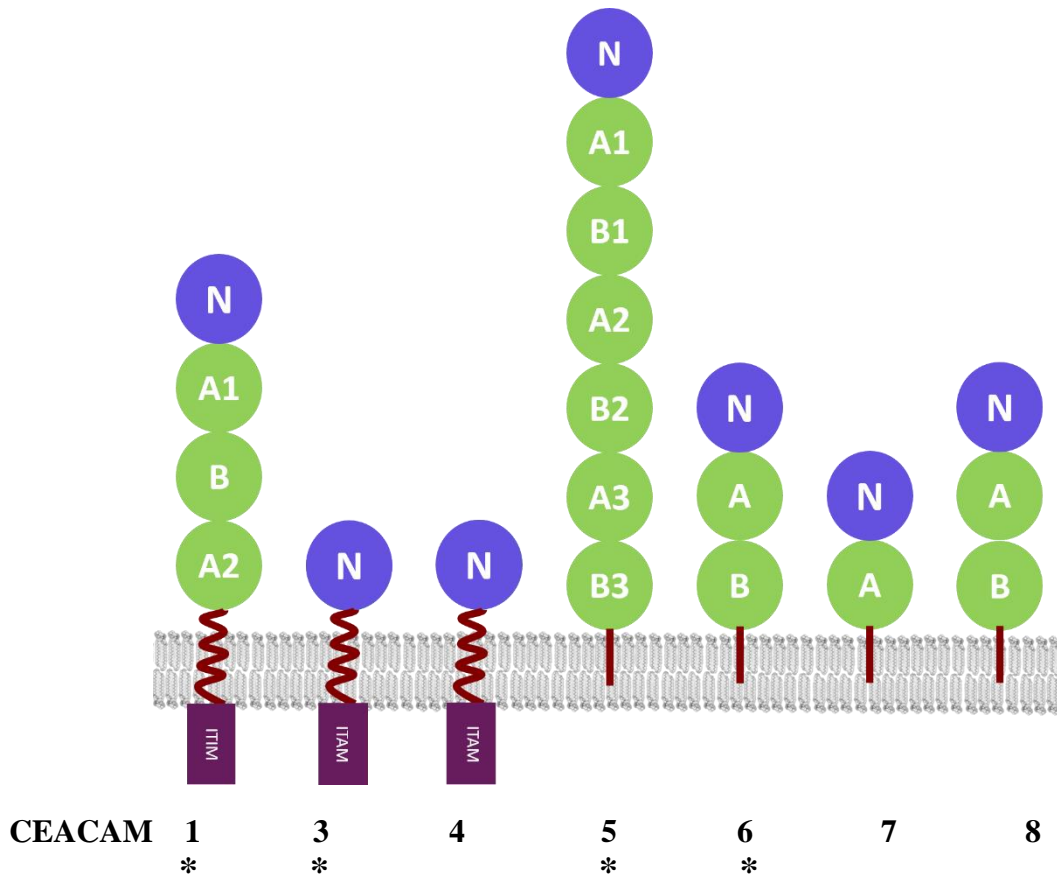


Figure 1.5: Domain organization of CEACAM receptors. The N-terminal IgV domain is shown in blue. The IgC domains are shown in green, and designated as either A or B. The element that links the CEACAM protein to the membrane is shown in red (either an α -helix for CCMs 1 and 3 (squiggly line), or a GPI element for CCMs 5-8 (bar)). Transmembrane domains are shown in purple, lipid bilayer is shown in gray. * indicates CEACAM receptors that interact with Opa_{CEA}.

1.3.2 CEACAMs act as a receptor for Opa proteins

CEACAM receptors mediate Gc and Nm colonization and engulfment by binding to Opa_{CEA} proteins. To date, only four CEACAMs have been shown to interact with Opa proteins: CEACAMs 1, 3, 5, and 6.⁴⁴ Opa_{CEA} proteins can interact specifically with only one of these CEACAMs, or with multiple different CEACAM variants.¹⁰ More specifically, the interaction occurs between the extracellular loops of Opa_{CEA} proteins, and the non-glycosylated face of the N-terminal IgV domain of the CEACAMs (NCCM, Fig. 1.6).^{13,45}

Ten CEACAM residues mediate binding to Opa proteins; of these important residues, only Tyr68 and Ile125 (residue numbers of CEACAM1, UniProt ID P13688-1, Fig 1.6, red) interact with all studied Opa variants, and are highly conserved on all CEACAMs.^{11,45} Of the other eight CEACAM residues involved in binding to Opa proteins, six are conserved between CEACAM1 and 3 (Fig. 1.6, orange). While the Opa-binding face of CEACAMs has been defined, in contrast, Opa sequence motifs that determine receptor specificity have not been identified.

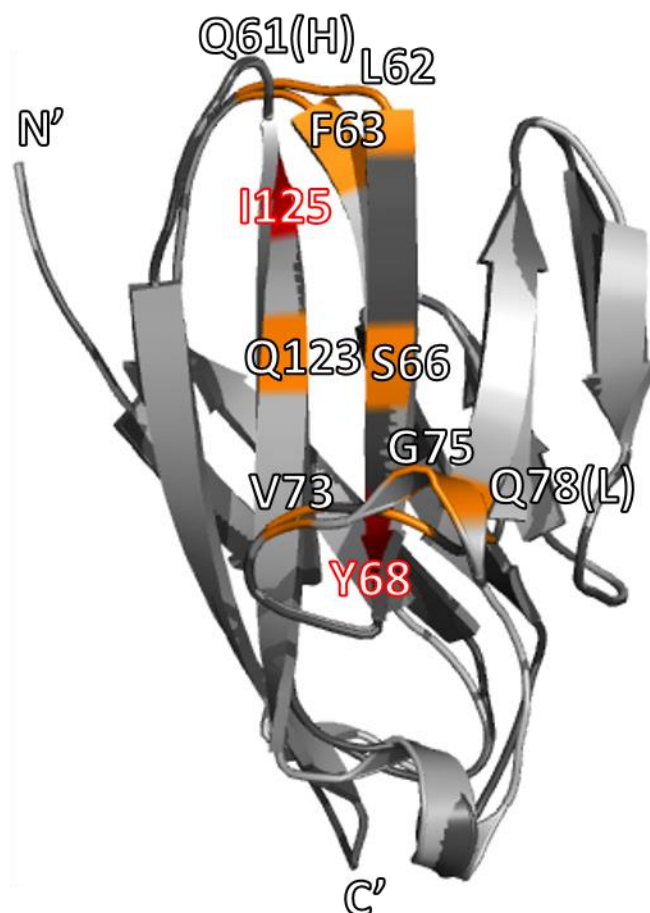


Figure 1.6: Cartoon representation of the N-terminal domain of human CEACAM1 (PDB ID 2GK2⁴⁵) and homology model of CEACAM3. NCCM1 is shown in light gray, and NCCM3 is depicted in dark gray (generated with SWISS-MODEL⁴⁶⁻⁴⁹). Residues that bind to all Opa_{CEA} proteins are colored red, while residues that only bind specific Opa variants are colored orange. All amino acids involved in Opa binding on NCCM1 are conserved in NCCM3, except Gln61 (His in NCCM3), and Gln78 (Leu in NCCM3).

1.3.3 CEACAM mediated engulfment of Opa expressing bacteria

Bacterial engulfment by the human host cell is a cellular response of CEACAMs interacting with Opa proteins. The mechanism of uptake differs depending on the CEACAM variant involved in the interaction.⁵⁰ The mechanism best understood so far is engulfment via CEACAM3, in which crown-like lamellipodial protrusions of actin filaments extend around individual or clustered bacteria, initiating engulfment.⁴⁴ The process is known to involve an immunoreceptor tyrosine-based activation motif (ITAM) on the CEACAM3 cytoplasmic domain (Fig 1.5, purple). CEACAM3 binding induces the phosphorylation of tyrosine residues in the ITAM by the Src family of kinases, and this phosphorylation triggers downstream activation of the small GTPases Rac and Cdc42, initiating F-actin assembly, leading to *Neisseria* internalization.^{24,51} CEACAM1 triggers a slightly different pathway for the engulfment of the bacteria, which involves pseudopods enveloping the bacteria.²⁴ Similar to CEACAM3, CEACAM1 has tyrosine residues on the cytoplasmic surface available for phosphorylation, but unlike CEACAM3 it does not induce a rearrangement of the actin.

Unlike CEACAM3 and CEACAM1, CEACAM5 and CEACAM6 are anchored to the membrane via a GPI element, and lack a cytoplasmic domain that can be phosphorylated. CEACAMs attached via a GPI element appear to engulf bacteria using a “zipper-like” mechanism, but the mechanism is not well understood.⁵² Binding of Opa to either CEACAM5 or CEACAM6 triggers progressive recruitment of adjacent receptors.⁴⁴ Eventually, this progression leads to complete envelopment of the bacterium by the cell membrane, which is “zippered” around the bacterium.⁴⁴

1.4 Liposomes are a useful tool for studying membrane proteins *in vitro*

In order to study purified membrane proteins *in vitro*, and for the protein to remain soluble, the protein must be placed in an environment similar to that of the lipid bilayer. This is accomplished by using a membrane mimic such as bicelles, nanodiscs, or more commonly detergent micelles or liposomes. Liposomes, vesicles formed by lipid bilayers, are common tools for biochemical research including cell tracking, non-viral gene transfer, and drug delivery.⁵³⁻⁵⁹ Liposomes can vary in size from ~50 nm to 1000 nm or more, and their size, composition, and surface properties can be easily manipulated for specific applications.⁶⁰ Liposomes may be designed to have multifunctional properties, including those that are pH or temperature sensitive, liposomes with steric stabilizing lipids (such as PEGylated lipids), and those with targeting ligands.⁶¹ The high level of customizability makes liposomes an excellent vehicle for studying membrane proteins.

Liposomes are particularly well suited for research involving membrane proteins involved in transport or catalysis, as the liposome is a closed vessel.⁶² There have been numerous studies of different membrane channels and transport systems using proteoliposomes (liposomes with a protein embedded in the bilayer), including KCNE1 (a modulator of voltage gated potassium channels like KCNQ1),⁶³ OEP24 (a general solute transporter in plants)⁶⁴, and bacteriorhodopsin (a proton pump).⁶⁵ Proteoliposomes are also used to study protein-assisted membrane fusion, such as via the SNARE complex,⁶⁶ influenza hemagglutinin for virus entry,⁶⁷ and fusion via the FAST (fusion-associated small transmembrane) family.⁶⁸ An additional benefit to using liposomes is that they can be tailored to closely mimic the phospholipid bilayer encapsulating most cells.

In addition to being utilized as research tools for membrane protein investigations, liposomes have medical applications as well. Liposomes are particularly well suited for use as delivery vehicles, as cargo can be encapsulated in the aqueous core, or embedded into the hydrophobic region of the bilayer. Additionally, they are biodegradable, biocompatible, and they reduce the degradation of the cargo, which leads to an increase in efficiency and a decrease in side effects.^{69,70} Liposomes are currently used for the delivery of many different therapeutics, including antifungals, antimicrobials, genes, and vaccines.^{69,71}

Despite these advantages, few liposomal systems are commercially available. These include liposomal doxorubicin (Myocet, Doxil, LipoDox, Thermodox) for the treatment several cancers, liposomal Amphotericin B (Ambisome) for the treatment of fungal infections, and liposomal vincristine (Marqibo) for the treatment of metastatic malignant uveal melanoma.⁷² Challenges facing liposomal delivery systems include intracellular delivery, cell-specific targeting, and shelf life.^{73,74}

1.5 Dissertation overview

Chapter 2 focuses on the Opa proteoliposome system, and techniques used to investigate the Opa – CEACAM interaction. A wide variety of lipids are commercially available, but not all lipids provide a suitable environment for membrane protein investigations. The rationale for the lipid composition used in *in vitro* Opa – CEACAM investigations is presented. Opa proteoliposomes were investigated for their interactions with CEACAMs using two different techniques, fluorescence polarization and biolayer interferometry. The theory behind these two techniques are described.

Chapter 3 presents the interaction of two CEACAM proteins (CEACAM1 and 3) with a variety of different Opa proteins in different environments. Binding was investigated using Opa expressing Gc, as well as Opa proteins reconstituted into liposomes. Pull-down assays and western blots were utilized to assess binding, and affinities were determined using fluorescence polarization. The results of this work demonstrate that our Opa proteoliposomes retain the ability to bind to CEACAM, validating their use in future *in vitro* studies.

Chapter 4 focuses on determining the binding affinities of a broader range of both CEACAM and Opa proteins. Understanding the kinetics and thermodynamics of a wide variety of Opa – CEACAM interactions may shed light on the Gc and Nm engulfment mechanism. Increasing our knowledge of bacterial engulfment can lead to better treatments for bacterial diseases, as well as the potential to exploit the internalization mechanism for targeted therapeutic delivery. A discussion of the calculated affinities and their biological relevance follows. Chapter 5 describes additional preliminary work and the future steps planned towards elucidating the molecular determinants of the Opa – CEACAM interaction.

1.6 References

1. CDC. Antibiotic Resistance Threats in the United States, 2013. (2013).
2. Ohnishi, M. et al. O3-S4.01 The new superbug *Neisseria gonorrhoeae* makes gonorrhoea untreatable?—first high-level ceftriaxone resistance worldwide and public health importance. *Sexually Transmitted Infections* **87**, A76 (2011).
3. Weber, D.J., Raasch, R. & Rutala, W.A. Nosocomial infections in the icu*: The growing importance of antibiotic-resistant pathogens. *Chest* **115**, 34S-41S (1999).
4. Meyer, T.F., Pohlner, J., and J. van Putten. Biology of the pathogenic *Neisseriae*. *Curr Topics Microbiol Immun* **192**, 283-317 (1994).
5. Association, T.N.M. Serogroup B Meningococcal Disease Outbreaks on U.S. College Campuses. Vol. 2016 (2016).
6. Todar, K. Todar's Online Textbook of Bacteriology. Vol. 2016 (Madison, Wisconsin, 2012).
7. Criss, A.K. & Seifert, H.S. A bacterial siren song: intimate interactions between *Neisseria* and neutrophils. *Nat Rev Micro* **10**, 178-190 (2012).
8. Hauck, C.R. & Meyer, T.F. 'Small' talk: Opa proteins as mediators of *Neisseria*–host-cell communication. *Curr Opin Microbiol* **6**, 43-49 (2003).
9. Sadarangani, M., Pollard, A.J. & Gray-Owen, S.D. Opa proteins and CEACAMs: pathways of immune engagement for pathogenic *Neisseria*. *FEMS Microbiology Reviews* **35**, 498-514 (2011).
10. Dehio, C., Gray-Owen, S.D. & Meyer, T.F. The role of neisserial Opa proteins in interactions with host cells. *TIM* **6**, 489-495 (1998).

11. Virji, M. et al. Critical determinants of host receptor targeting by *Neisseria meningitidis* and *Neisseria gonorrhoeae* : identification of Opa adhesiotopes on the N-domain of CD66 molecules. *Mol Microbiol* **34**, 538-551 (1999).
12. Ball, L.M. & Criss, A.K. Constitutively Opa-expressing and Opa-deficient *neisseria gonorrhoeae* strains differentially stimulate and survive exposure to human neutrophils. *J Bacteriol* **195**, 2982-90 (2013).
13. Bos, M.P., Grunert, F. & Belland, R.J. Differential recognition of members of the carcinoembryonic antigen family by Opa variants of *Neisseria gonorrhoeae*. *Infect Immun* **65**, 2353-61 (1997).
14. Muenzner, P. et al. Carcinoembryonic Antigen Family Receptor Specificity of *Neisseria meningitidis* Opa Variants Influences Adherence to and Invasion of Proinflammatory Cytokine-Activated Endothelial Cells. *Infect Immun* **68**, 3601-3607 (2000).
15. Bilek, N., Ison, C.A. & Spratt, B.G. Relative Contributions of Recombination and Mutation to the Diversification of the opa Gene Repertoire of *Neisseria gonorrhoeae*. *J Bacteriol* **191**, 1878-1890 (2009).
16. Merz, A.J. & So, M. Interactions of Pathogenic *Neisseriae* with Epithelial Cell Membranes. *Annu Rev Cell Dev Biol* **16**, 423-457 (2000).
17. Fox, D.A. et al. Structure of the Neisserial Outer Membrane Protein Opa60: Loop Flexibility Essential to Receptor Recognition and Bacterial Engulfment. *JACS* **136**, 9938-9946 (2014).

18. Malorny, B., Morelli, G., Kusecek, B., Kolberg, J. & Achtman, M. Sequence Diversity, Predicted Two-Dimensional Protein Structure, and Epitope Mapping of Neisserial Opa Proteins. *J Bacteriol* **180**, 1323-1330 (1998).
19. Bos, M.P., Kao, D., Hogan, D.M., Grant, C.C.R. & Belland, R.J. Carcinoembryonic Antigen Family Receptor Recognition by Gonococcal Opa Proteins Requires Distinct Combinations of Hypervariable Opa Protein Domains. *Infect Immun* **70**, 1715-1723 (2002).
20. Bos, M.P., Kao, D., Hogan, D.M., Grant, C.C. & Belland, R.J. Carcinoembryonic antigen family receptor recognition by gonococcal Opa proteins requires distinct combinations of hypervariable Opa protein domains. *Infect Immun* **70**, 1715-23 (2002).
21. Swanson, J.O., Barrera, J. Sola, and Boslego, J. Expression of outer membrane protein II by gonococci in experimental gonorrhea. *J Exp Med* **168**, 2121-2129 (1988).
22. Bhat, K.S. et al. The opacity proteins of *Neisseria gonorrhoeae* strain MS11 are encoded by a family of 11 complete genes. *Mol Microbiol* **5**, 1889-1901 (1991).
23. Kupsch, E.M., Knepper, B., Kuroki, T., Heuer, I. & Meyer, T.F. Variable opacity (Opa) outer membrane proteins account for the cell tropisms displayed by *Neisseria gonorrhoeae* for human leukocytes and epithelial cells. *EMBO J* **12**, 641-650 (1993).
24. Billker, O. et al. Distinct mechanisms of internalization of *Neisseria gonorrhoeae* by members of the CEACAM receptor family involving Rac1- and Cdc42-dependent and -independent pathways. *EMBO J* **21**, 560-571 (2002).
25. Abdul-Wahid, A., Huang, E.H.B., Cydzik, M., Bolewska-Pedyczak, E. & Gariépy, J. The carcinoembryonic antigen IgV-like N domain plays a critical role in the implantation of metastatic tumor cells. *Mol Oncol* **8**, 337-350 (2014).

26. Lu, R., Kujawski, M., Pan, H. & Shively, J.E. Tumor angiogenesis mediated by myeloid cells is negatively regulated by CEACAM-1. *Cancer Res* **72**, 2239-2250 (2012).
27. Kuespert, K., Pils, S. & Hauck, C.R. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol* **18**, 565-571 (2006).
28. Zebhauser, R. et al. Identification of a novel group of evolutionarily conserved members within the rapidly diverging murine Cea family. *Genomics* **86**, 566-580 (2005).
29. Benchimol, S. et al. Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. *Cell* **57**, 327-334 (1989).
30. Ordoñez, C., Screaton, R.A., Ilantzis, C. & Stanners, C.P. Human Carcinoembryonic Antigen Functions as a General Inhibitor of Anoikis. *Cancer Res* **60**, 3419-3424 (2000).
31. Ergün, S. et al. CEA-Related Cell Adhesion Molecule 1: A Potent Angiogenic Factor and a Major Effector of Vascular Endothelial Growth Factor. *Mol Cell* **5**, 311-320 (2000).
32. Poy, M.N. et al. CEACAM1 regulates insulin clearance in liver. *Nat Genet* **30**, 270-276 (2002).
33. Leung, N. et al. Deletion of the carcinoembryonic antigen-related cell adhesion molecule 1 (Ceacam1) gene contributes to colon tumor progression in a murine model of carcinogenesis. *Oncogene* **25**, 5527-5536 (2006).
34. Hauck, C.R., Agerer, F., Muenzner, P. & Schmitter, T. Cellular adhesion molecules as targets for bacterial infection. *Eur J Cell Biol* **85**, 235-242 (2006).

35. Johnson, M.B. et al. Opa+ *Neisseria gonorrhoeae* exhibits reduced survival in human neutrophils via Src family kinase-mediated bacterial trafficking into mature phagolysosomes. *Cell Microbiol* **17**, 648-665 (2015).
36. Nagel, G. et al. Genomic organization, splice variants and expression of CGM1, a SD66-related member of the carcinoembryonic antigen gene family. *Eur J Biochem* **214**, 27-35 (1993).
37. Chen, T. & Gotschlich, E.C. CGM1a antigen of neutrophils, a receptor of gonococcal opacity proteins. *Proc Nat Acad Sci U S A* **93**, 14851-14856 (1996).
38. Kuroki, M. et al. Molecular cloning of nonspecific cross-reacting antigens in human granulocytes. *J Biol Chem* **266**, 11810-11817 (1991).
39. Thompson, J.A. Molecular Cloning and Expression of Carcinoembryonic Antigen Gene Family Members. *Tumor Biol* **16**, 10-16 (1995).
40. Schölzel, S. et al. Carcinoembryonic Antigen Family Members CEACAM6 and CEACAM7 Are Differentially Expressed in Normal Tissues and Oppositely Deregulated in Hyperplastic Colorectal Polyps and Early Adenomas. *Am J Pathol* **156**, 595-605 (2000).
41. Buchegger, F., Schreyer, M., Carrel, S. & Mach, J.-P. Monoclonal antibodies identify a cea crossreacting antigen of 95 kD (NCA-95) distinct in antigenicity and tissue distribution from the previously described NCA of 55 kD. *Int J Cancer* **33**, 643-649 (1984).
42. Wang, J., Gray-Owen, S.D., Knorre, A., Meyer, T.F. & Dehio, C. Opa binding to cellular CD66 receptors mediates the transcellular traversal of *Neisseria gonorrhoeae* across polarized T84 epithelial cell monolayers. *Mol Microbiol* **30**, 657-71 (1998).

43. Simons, K. & Toomre, D. Lipid rafts and signal transduction. *Nat Rev Mol Cell Biol* **1**, 31-9 (2000).
44. McCaw, S.E., Liao, E.H. & Gray-Owen, S.D. Engulfment of *Neisseria gonorrhoeae*: Revealing Distinct Processes of Bacterial Entry by Individual Carcinoembryonic Antigen-Related Cellular Adhesion Molecule Family Receptors. *Infect Immun* **72**, 2742-2752 (2004).
45. Fedarovich, A., Tomberg, J., Nicholas, R.A. & Davies, C. Structure of the N-terminal domain of human CEACAM1: binding target of the opacity proteins during invasion of *Neisseria meningitidis* and *N. gonorrhoeae*. *Acta Crystallogr Sect D-Biol Crystallogr* **62**, 971-979 (2006).
46. Guex, N., Peitsch, M.C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30 Suppl 1**, S162-73 (2009).
47. Biasini, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252-8 (2014).
48. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201 (2006).
49. Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res* **37**, D387-92 (2009).
50. Chen, T., Grunert, F., Medina-Marino, A. & Gotschlich, E.C. Several Carcinoembryonic Antigens (CD66) Serve as Receptors for Gonococcal Opacity Proteins. *J Exp Med* **185**, 1557-1564 (1997).

51. Booth, J.W. et al. Phosphatidylinositol 3-Kinases in Carcinoembryonic Antigen-related Cellular Adhesion Molecule-mediated Internalization of *Neisseria gonorrhoeae*. *J Biol Chem* **278**, 14037-14045 (2003).
52. Swanson, J.A. & Baer, S.C. Phagocytosis by zippers and triggers. *Trends in Cell Biol* **5**, 89-93 (1995).
53. Csiszár, A. et al. Novel Fusogenic Liposomes for Fluorescent Cell Labeling and Membrane Modification. *Bioconjugate Chem* **21**, 537-543 (2010).
54. Gopalakrishnan, G. et al. Multifunctional Lipid/Quantum Dot Hybrid Nanocontainers for Controlled Targeting of Live Cells. *Angew Chem Int Ed* **45**, 5478-5483 (2006).
55. Seo, J.W., Zhang, H., Kukis, D.L., Meares, C.F. & Ferrara, K.W. A Novel Method to Label Preformed Liposomes with ⁶⁴Cu for Positron Emission Tomography (PET) Imaging. *Bioconjugate Chem* **19**, 2577-2584 (2008).
56. Chesnoy, S.a.H., L. Structure and Function of Lipid-DNA Complexes for Gene Delivery. *Annu Rev Biophys Biomol Struct* **29**, 27-47 (2000).
57. Kawakami, S., Fumoto, S., Nishikawa, M., Yamashita, F. & Hashida, M. In Vivo Gene Delivery to the Liver Using Novel Galactosylated Cationic Liposomes. *Pharmaceut Res* **17**, 306-313 (2000).
58. Narang, A.S., Thoma, L., Miller, D.D. & Mahato, R.I. Cationic Lipids with Increased DNA Binding Affinity for Nonviral Gene Transfer in Dividing and Nondividing Cells. *Bioconjugate Chem* **16**, 156-168 (2005).
59. Sawant, R.M. et al. "SMART" Drug Delivery Systems: Double-Targeted pH-Responsive Pharmaceutical Nanocarriers. *Bioconjugate Chem* **17**, 943-949 (2006).

60. Rigaud, J.-L. & Lévy, D. Reconstitution of Membrane Proteins into Liposomes. in *Method Enzymol*, Vol. Volume 372 65-86 (Academic Press, 2003).
61. Petros, R.A. & DeSimone, J.M. Strategies in the design of nanoparticles for therapeutic applications. *Nat Rev Drug Discov* **9**, 615-627 (2010).
62. Rigaud, J.-L. Membrane proteins: functional and structural studies using reconstituted proteoliposomes and 2-D crystals. *Braz J Med Biol Res* **35**, 753-766 (2002).
63. Sahu, I.D. et al. Structural Investigation of the Transmembrane Domain of KCNE1 in Proteoliposomes. *Biochemistry* **53**, 6392-6401 (2014).
64. Liguori, L., Blesneac, I., Madern, D., Vivaudou, M. & Lenormand, J.-L. Single-step production of functional OEP24 proteoliposomes. *Protein Express Purif* **69**, 106-111 (2010).
65. Sandén, T., Salomonsson, L., Brzezinski, P. & Widengren, J. Surface-coupled proton exchange of a membrane-bound proton acceptor. *Proc Nat Acad Sci U S A* **107**, 4129-4134 (2010).
66. Hernandez, J.M. et al. Membrane Fusion Intermediates via Directional and Full Assembly of the SNARE Complex. *Science* **336**, 1581-1584 (2012).
67. Lai, A.L. & Tamm, L.K. Shallow Boomerang-shaped Influenza Hemagglutinin G13A Mutant Structure Promotes Leaky Membrane Fusion. *J Biol Chem* **285**, 37467-37475 (2010).
68. Top, D. et al. Liposome reconstitution of a minimal protein-mediated membrane fusion machine. *EMBO J* **24**, 2980-2988 (2005).

69. Al-Jamal, W.T. & Kostarelos, K. Liposomes: From a Clinically Established Drug Delivery System to a Nanoparticle Platform for Theranostic Nanomedicine. *Acc Chem Res* **44**, 1094-1104 (2011).
70. Torchilin, V.P. Recent advances with liposomes as pharmaceutical carriers. *Nat Rev Drug Discov* **4**, 145-160 (2005).
71. Lasic, D., and Papahadjopolous, D. *Medical Applications of Liposomes*, (Elsevier Science, Amsterdam, 1998).
72. FDA. FDA Approved Drug Products. Vol. 2016 (U.S. Food and Drug Administration, Silver Springs, MD).
73. Torchilin, V.P., Rammohan, R., Weissig, V. & Levchenko, T.S. TAT peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors. *Proc Nat Acad Sci U S A* **98**, 8786-8791 (2001).
74. Dutta, R.C. Drug Carriers in Pharmaceutical Design: Promises and Progress. *Curr Pharm Des* **13**, 761-769 (2007).

2. Quantitative methods used to investigate the Opa – CEACAM interaction

2.1 Overview

To date, all investigations of Opa protein interactions with receptors have been performed *in vivo*, either in Gc or *E. coli*. Investigating Opa proteins *in vitro* allows for quantitative analysis of receptor selectivity and affinity, as well as the determination of specific cellular outcomes mediated solely by Opa proteins. In order to investigate membrane proteins, such as the Neisserial Opa proteins, a membrane mimic must be utilized to maintain protein solubility. Liposomes are one such membrane-like structure commonly used. However, there is a vast collection of lipid molecules available to make liposomes, and not all combinations allow proteins to be stable, or more importantly, functional. This chapter outlines the strategy behind the selection of lipid composition for *in vitro* investigations of Opa protein – CEACAM interactions. Additionally, this chapter will provide theory behind the several different techniques that were utilized to investigate binding between these two proteins, discussed in chapters 3 and 4.

2.2 Selection of lipid composition for *in vitro* investigations of Opa – CEACAM interactions

2.2.1 Membrane proteins in research

Proteins are grouped into two classes: soluble proteins and membrane proteins. Soluble proteins are in an aqueous environment, while membrane proteins contain regions that are embedded into lipid bilayers. Membrane proteins are responsible for the majority of cell interactions with their environment. Additionally, membrane proteins play key roles in ion transport into and out of cells, energy transduction, catalysis, signal recognition, and

transduction.¹ Because of this wide variety of biological functions, membrane proteins are often the target for pharmaceuticals. Approximately 50% of commercial drugs target membrane proteins.² In addition to being critical to many biological functions, membrane proteins are ubiquitous; approximately 30% of the coding genome encodes membrane proteins.³

Despite their prevalence and importance, membrane protein research has lagged compared to soluble proteins. Fewer than 1% of the unique structures deposited in the Protein Data Bank (PDB) belong to membrane proteins.¹ The first soluble protein structure was determined in 1960, which was 25 years before the first membrane protein structure in 1985.⁴ Despite recent technological advances, the rate of membrane protein structure determination is slower than that of soluble proteins 25 years earlier (Figure 2.1) due to difficulties in membrane protein recombinant expression, solubilization, and folding.

Computational analysis of the *E. coli* genome identified over 800 membrane proteins (~100 in the outer membrane, the rest residing in the inner membrane).⁵ Of the membrane proteins residing in the inner membrane, approximately one third have an unknown function.⁵ Of those that are classified into a functional class, little more is known than the general category (i.e. involved in signaling, transport, channel, etc.).⁵ For the outer membrane β -barrel proteins, the function remains unknown for one third to one half.⁵

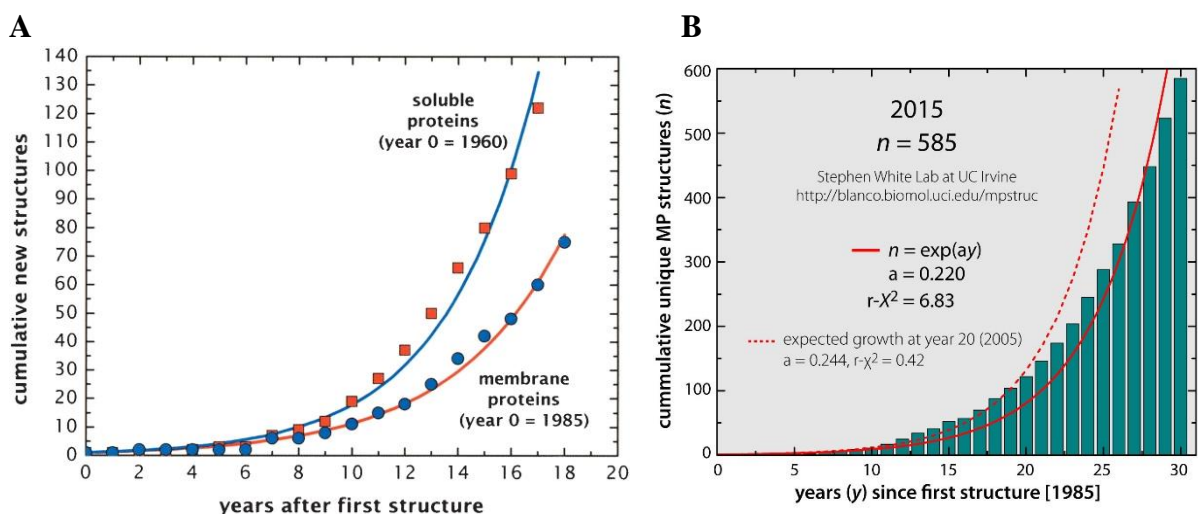


Figure 2.1. Progress in membrane protein structure determination. In 2004, the number of structures of both soluble proteins (red squares) and membrane proteins (blue circles) was plotted for the 20 years following the first structure determination (A). Membrane protein structure determination, while growing exponentially, is lagging behind soluble protein structure determination.⁶ An updated 2015 analysis of membrane protein structure determination indicates that the exponential growth predicted in 2005 (dotted line) exceeded the actual progress (solid line), despite the major advances facilitating GPCR structures (B).⁷

2.2.2 Utilizing liposomes to investigate membrane proteins

Interactions of Opa proteins with receptors have been investigated thus far in both Gc and *E. coli*. However, *in vitro* experiments must be utilized to investigate Opa-induced cellular responses, as well as to quantify the interactions of Opa proteins with receptors. Because Opas are membrane proteins, they must be folded into a membrane mimic to be soluble and functional. Biological investigations can only be undertaken after assuring that Opa proteins are fully folded and reconstituted into a membrane-like environment.

The forces involved in membrane protein stability are not well understood. As such, most investigators empirically screen different lipids or detergents and solution conditions to produce stable, folded membrane proteins for *in vitro* investigations. Recombinant folding directly into liposomes has been demonstrated for a small number of β -barrel proteins from a urea denatured state, including OmpA from *E. coli*,⁸⁻¹¹ OmpF from *E. coli*,¹² OmpX from *E. coli*,¹³ and VDAC from mitochondrial outer membranes of humans.¹⁴ A systematic study which optimized the folding of nine outer membrane proteins (OmpLa, PagP, OmpX, OmpT, Omp85, OmpW, FadL, OmpA, and OmpF) found widely varying efficiencies for folding into the same bilayer.¹⁵ This study demonstrates that systematic screening of folding conditions for reconstitution into liposomes is necessary to produce stable, fully folded proteins to be used for further investigations.

As mentioned in Chapter 1, liposomes are spherical vesicles composed of lipid bilayers that can be used to study membrane proteins *in vitro*. Lipids have a lower critical micelle concentration (CMC, defined as the amount of surfactant required to form ordered structures) than detergents, which allows biological experiments to be performed with a lower amphiphile background. This is advantageous, as free amphiphiles may interfere with binding experiments

through non-specific interactions, or whole-cell experiments by the disrupting the cell membrane. The bilayer environment of a liposome is more similar to native cell membrane bilayers than detergent micelle environments, which is also advantageous in studying biological functions of proteins. Additionally, Liposomes can be used in whole cell experiments, providing a sample that can be used for both *in vitro* and *in vivo* experiments.

2.2.3 Developing a liposomal system for Opa proteins

The folding of Opa proteins into liposomes was described by Dewald *et al* in 2011.¹⁶ The conditions in which Opa proteins fold most effectively are not conducive to biological investigations. A highly basic (pH 12) buffer containing 4 M urea is needed, with short chain (diC10PC) lipids.¹⁶ Opa folding occurs over a period of three days with heat (37 °C).¹⁶ After this lengthy folding time, the integrity of the structure of lipids surrounding the Opa proteins is compromised, due to degradation of the lipid molecules.¹⁷ Liposomes composed of short chain lipids are not as stable as those formed by longer chain lipids, as the short hydrocarbon tails have less van der Waals forces between them.¹⁸

Such folding conditions yield an environment that is not appropriate for further biological investigations of Opa proteins. Additionally, the large amount of lipid used for folding Opa does not allow Opa to pellet upon ultra-centrifugation (the lipid:protein ratio is too high), so empty liposomes could not be used for pull-down assays (described in chapter 3). As such, after folding, Opa proteins were transferred to more stable liposomes composed of the longer carbon chain lipids DMPC and DMPG (14 carbon tails), along with cholesterol for rigidity, via pelleting and resuspension, using a sonication method developed and described by

Alison Dewald.¹⁷ This lipid composition produces SUVs which are stable over a period of several days, in a buffer that is more conducive for biological experiments.

However, there are a few downfalls of using SUVs. SUVs are less stable than their larger counterparts. While LUVs (large unilamellar vesicles) will form spontaneously over time, the formation of SUVs requires energy unless charged lipids are present.¹⁹ Smaller vesicles are more strained than their large counterparts, as they have a higher curvature.²⁰ Higher membrane curvature promotes the fusion of lipid membranes.²¹ Additionally, small vesicles typically exhibit a higher amount of non-specific interactions.²²

One way to combat these negative attributes of SUVs is through the addition of polyethylene glycol (PEG)-ylated lipids into the liposome. PEGylated lipids have been shown to decrease the non-specific interactions with the liposomes, as well as to prevent the fusion of vesicles.²³⁻²⁵ The addition of PEGylated lipids to the liposomes not only aids in preventing non-specific interactions, but may also provide a better membrane mimic for Opa proteins. The outer membrane of *Neisseria* bacteria contains lipo-oligosaccharides (LOS), a shortened form of the lipopolysaccharides (LPS) found on other bacteria.²⁶ As LOS is the primary component of the *Neisseria* outer membrane, it is likely that the LOS molecules interact with Opa proteins via the basic residues on the extracellular loops,²⁷ which modulates the mobility of the extracellular loops. Incorporating PEGylated lipids into Opa proteoliposome may mimic the spatial restrictions that LOS has on the Opa protein extracellular loops, but does not mimic the electrostatic interactions of LOS with the Opa loops.

2.2.4 Selection of PEGylated lipids

PEGylated lipids have been utilized in liposomes for drug delivery since the early 1990's.²⁸ Previously, a major limitation to therapeutic liposomal delivery systems was their rapid clearance by macrophages before delivering their cargo.²⁴ Woodle *et al* reported that the addition of PEGylated phosphoethanolamine (PE) lipids prolonged circulation of the liposomes and decreased their uptake by macrophages in mice.²⁸ The increased circulation time of PEGylated liposomes is likely due to reduced interactions with cell adhesion proteins, however a conclusive link has not been established.²⁹⁻³¹

Most PEGylated lipids consist of PEG polymerized to a PE head group with varying lipid tail lengths. A variety of lipids with different PEG polymer molecular weights are commercially available and used in research.^{23,24} The behavior of the PEGylated liposome is dependent on both the amount of PEG present and the molecular weight of the PEG polymer.³¹ Many biological experiments and liposomal therapeutics utilize 5 mol% PEGylation in their liposomal system.^{24,31,32} The molecular weight of incorporated PEG-polymerized lipids typically ranges between 750 – 5000 Daltons (PEG750 – PEG5000).^{23,24,31}

Needham *et al* incorporated PEG1900 into their liposomes, and determined that the polymer extended approximately 50 Å from the surface of the lipids.³³ However, others have reported that PEG polymers form various structures on the liposomes, including “mushrooms,” “brushes,” and “pancakes” (Figure 2.2).^{31,34} Du *et al* report that no PEG structure model alone is able to explain protein adsorption.²³ As such, it is logical to conclude that PEG polymers on the surface of the liposome are in various conformations and extend various distances past the bilayer. The extracellular loops of Opa₆₀ extend approximately 20 – 30 Å above the barrel and membrane. The loops must be accessible for receptor interactions, but may also need spatial

restrictions to mimic LOS of *Neisseria*. Therefore, PEG1000 was selected to add to the lipid composition. PEG1000 may spatially restrict the motion of the Opa extracellular loops, while potentially decreasing non-specific interactions with the liposomes. A final liposome composition of DMPC (63 mol%), DMPG (16 mol%), cholesterol (16 mol%), and DMPE-PEG1000 (5 mol%) was chosen to be used for further biological investigation of the Opa – CEACAM interactions *in vitro*.

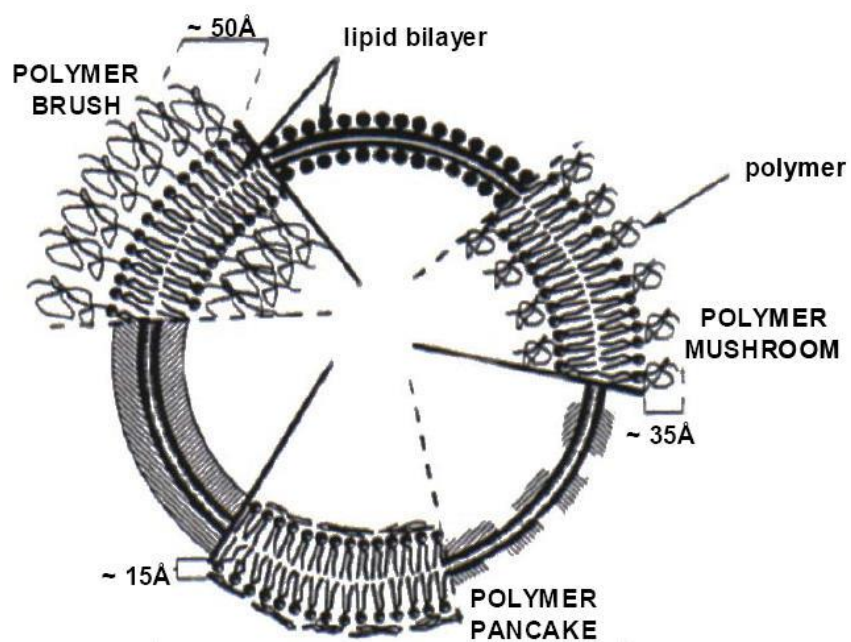


Figure 2.2. Possible conformations of PEG attached to a membrane structure. *Reprinted with permission.*³¹ The PEG polymer is capable of forming various structures on the surface of the liposome, including the brush, mushroom, and pancake.

2.2.5 Trypsin cleavage of extracellular Opa protein loops

Following Opa reconstitution into liposomes and exchange into a more stable and biologically appropriate liposome system, the orientation of Opa proteins within the liposome must be assessed before functional assays can be performed. When most β -barrel membrane proteins fold *in vivo*, they are inserted into the outer membrane and folded by a complex of proteins referred to collectively as the β -barrel assembly machinery (BAM complex).³⁵ The BAM complex ensures that the proteins are inserted into the membrane in the correct orientation: the extracellular loops on the exterior of the cell, and the periplasmic turns remaining in the periplasm.

Since reconstitution into liposomes is spontaneous folding without the use of chaperones, it is necessary to assess the orientation once the Opa proteins are folded in liposomes. If the Opa proteins are inserted into the liposomes with the extracellular loops facing the interior, they will not be available to bind receptors. Therefore, to quantify Opa – receptor interactions, the percentage of Opa proteins in an orientation that allows the loops to interact with receptors must be determined.

Previously, the orientation of OmpA after folding into liposomes was determined via an enzymatic trypsin assay.⁸ The periplasmic domain of OmpA consists of a 24 kDa segment that is sensitive to cleavage by trypsin. Trypsin is a serine protease that cleaves peptide chains at the carboxyl side of the positive amino acids (lysine and arginine), except when followed by a proline residue. Surrey *et al* incubated OmpA proteoliposomes with trypsin and assessed cleavage via SDS-PAGE. They found that after folding OmpA directly into liposomes, there is almost complete cleavage of the periplasmic domain, indicating that the extracellular portion of OmpA is located in the interior of the liposomes.⁸

The orientation of Opa proteins in their final lipid mixture was assessed in a similar manner. NMR and SDS-PAGE analysis of trypsin-cleaved Opa proteins folded into detergent micelles indicates that upon incubation with trypsin, the extracellular loops are removed but the barrel domain stays intact and folded.³⁶ Therefore, the percentage of Opa proteins whose loops are not cleaved after incubation with trypsin would be indicative of the amount of Opa proteins inserted facing the interior of the liposome that would not be available for receptor interactions. SDS-PAGE analysis of trypsin treated Opa proteoliposomes indicates that approximately half of the Opa proteins are oriented with their extracellular loops facing outwards, and able to interact with receptors, while half are not (Figure 2.3). After 24 hours of incubation with trypsin, the Opa proteoliposomes begin to show a higher amount of cleavage product, which is most likely due to fusion of the SUVs.

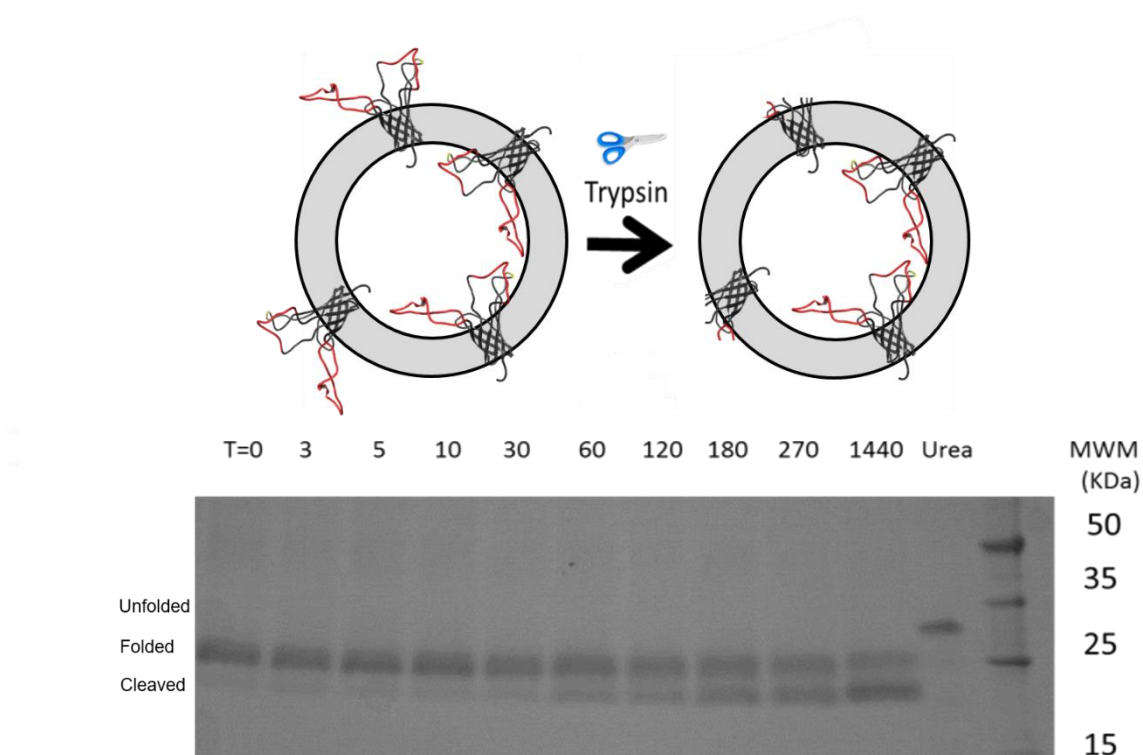
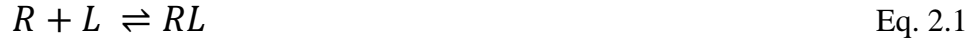


Figure 2.3. Trypsin cleavage of Opa proteoliposomes. Opa proteoliposomes were incubated with trypsin and analyzed using SDS-PAGE. Time (T) is minutes after trypsin addition. The lipid lane (T = 0) shows folded, intact Opa proteins before addition of trypsin, and the urea lane shows unfolded Opa protein. The cleaved product is the intact β -barrel domain of the Opa protein, without the extracellular loops.

2.3 Understanding thermodynamics of protein – ligand interactions

For a protein – ligand interaction in which one receptor molecule (R) binds one ligand (L), the binding equilibrium is defined as:



with a dissociation constant (K_D) equal to:

$$K_D = \frac{[RL]}{[R][L]} \quad \text{Eq. 2.2}$$

Consider Y to be the degree of saturation or fraction bound (ratio of moles of bound ligand over moles of total binding sites):

$$Y = \frac{[bound\ ligand]}{[total\ binding\ sites]} \quad \text{Eq. 2.3}$$

With this equation, the concentration of the receptor – ligand complex ([RL]) may be equated to $Y[R]_t$, where $[R]_t$ is the total concentration of receptor in solution. The concentration of unbound receptor $[R]_u$ becomes:

$$[R]_u = (1 - Y)[R]_t \quad \text{Eq. 2.4}$$

Equation 2.2 can be rewritten as:

$$K_D = \frac{Y[R]_t}{[L](1-Y)[R]_t} \quad \text{Eq. 2.5}$$

Which, when simplified, can be rearranged to yield:

$$Y = \frac{K_D[L]}{1 + K_D[L]} \quad \text{Eq. 2.6}$$

Equation 2.6 is the Langmuir binding isotherm equation, which makes four assumptions: (i) each binding site can be occupied by only one ligand, (ii) each binding site has an identical affinity for the ligand, (iii) the solvent does not participate, and (iv) the system is sufficiently dilute that extraneous interactions do not interfere with binding.³⁷ In the case of Opa –

CEACAM interactions, the Langmuir binding equation is applied to proteins which have a single binding site.

2.4 Fluorescence polarization (FP) is well suited to study the Opa – CEACAM interaction

2.4.1 Underlying principles of FP

Quantifying interactions between biomolecules is essential to understanding biological function. Fluorescence polarization (FP) is one method frequently used to investigate protein – ligand interactions *in vitro*. Fluorescence polarization can be broken down into two main concepts: fluorescence and polarization (or anisotropy).

The Jablonski energy diagram describes the principle of fluorescence (Figure 2.4). To have fluorescence, a molecule must absorb a photon, exciting an electron to a higher energy state. The electron loses some of the energy in the form of heat and then releases a photon, relaxing to a lower energy state after a short period of time (ns) . How long this relaxation takes is termed the fluorescence lifetime. The emitted light is lower in energy than the light used to excite the molecule, and so the emitted light has a longer wavelength, referred to as the Stokes shift.

Polarization of light results in waves that are oscillating in a single direction. When polarized light is applied to fluorophores that are randomly oriented in solution, the fluorophores most likely to be excited are those oriented within a particular range of angles to the applied polarization.³⁸ If those fluorophores do not move, the emitted light will also be polarized within a particular range of angles to the applied light.³⁸ However, all of the molecules are undergoing Brownian motion, and will be tumbling freely in solution. As the molecules rotate in solution, the emitted light will no longer be the polarized. Therefore,

polarization measurements are used to detect the rotation of a molecule in a particular environment. The fluorescence polarization can be defined quantitatively by the following equation:

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \quad \text{Eq. 2.7}$$

where P is polarization, I_{\parallel} is the intensity of light emitted parallel to the excitation light, and I_{\perp} is the intensity of light perpendicular to the excitation light.³⁹

In the FP experimental set-up, the smaller species are fluorescently labeled and held at a constant, low concentration.⁴⁰ The larger, unlabeled molecule will be titrated in to have a maximal change in fluorescence polarization between the free ligand and the bound state. The small, fluorescently labeled species will be tumbling rapidly in solution, and will emit depolarized light (Figure 2.5, top). Upon interaction with the larger species, the complex will tumble at a slower rate, and the emitted light from the bound complex will be polarized (Figure 2.5, bottom). The differences in polarization between the bound and unbound states correlate to the fraction of the small ligand that is bound to the receptor. An association curve can be generated by plotting the fraction bound vs. the concentration of receptor, using the following equation:

$$P = F_f P_f + F_b P_b \quad \text{Eq. 2.8}$$

where P is the observed polarization value, F_f is the fraction of fluorescent ligand free, F_b is the fraction of fluorescent ligand bound, P_f is the polarization of the free fluorescent ligand, and P_b is the polarization of the bound fluorescent ligand.

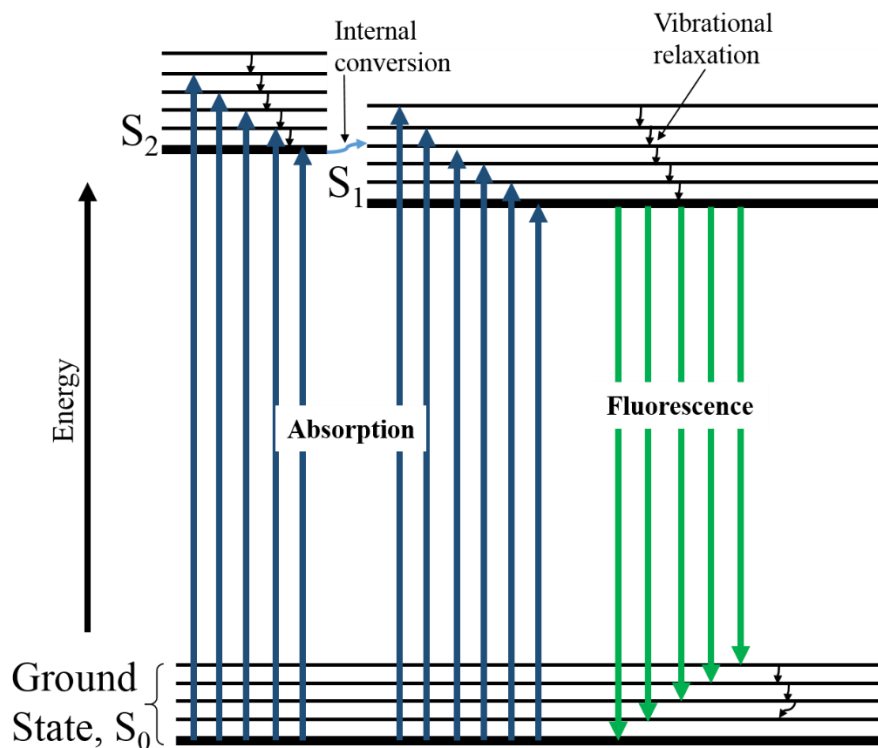


Figure 2.4. The Jablonski energy diagram explains the principles of fluorescence. A molecule is excited from the ground state by energy in the form of photons, and relaxes down to a lower energy state. Fluorescence is observed when a photon is emitted to allow the molecule to return to the ground state.

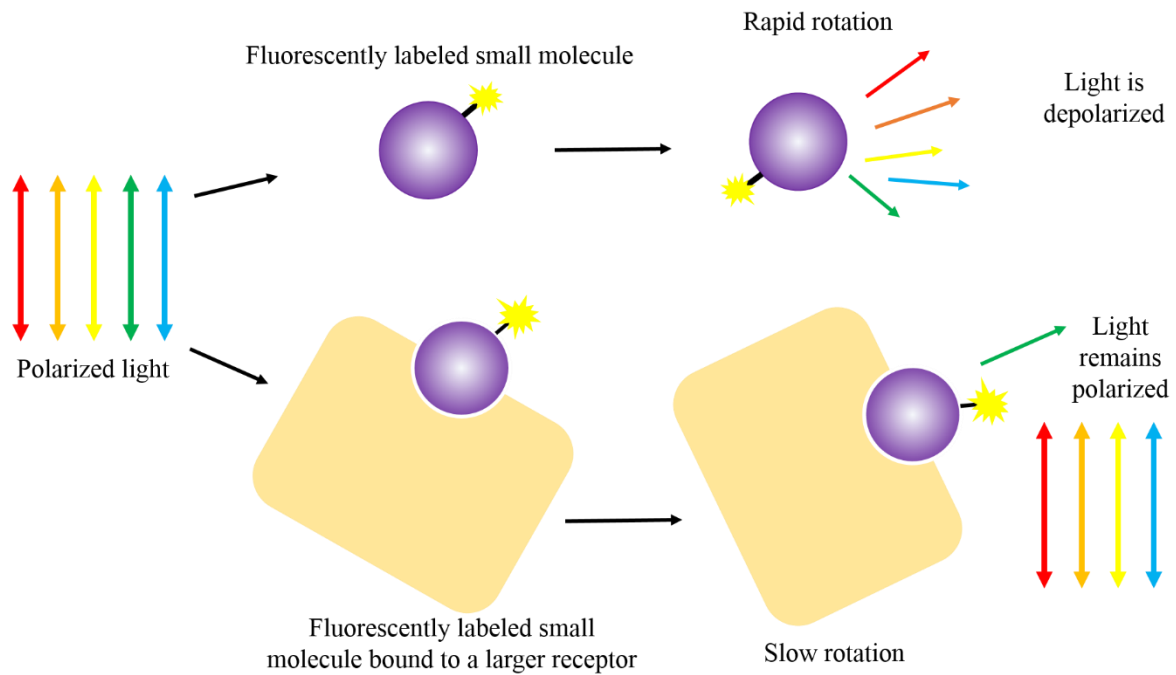


Figure 2.5. The principles of fluorescence polarization. A fluorophore attached to a ligand is excited by plane polarized light. Small molecules will tumble rapidly in solution, and the emitted light will be depolarized. Larger molecules will tumble slower in solution, and the emitted light will retain some of the polarization.

2.4.2 Opa – CEACAM interactions can be assayed using FP

Typically, FP is used to measure affinities between proteins and small molecules. However, since the depolarization of emitted light is dependent on the tumbling of the fluorescently labeled molecule, larger molecules may be used with fluorophores that have a longer fluorescent lifetime. The fluorophore chosen for our Opa – CEACAM system is 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid, disodium salt (AMS). According to the manufacturer, this dye has a fluorescent lifetime of approximately 10 ns. The molecular tumbling of a molecule in solution is defined by the following equations:

$$\tau_c = \frac{kT}{8\pi\eta a^3} \quad \text{Eq. 2.9}$$

$$\tau_c = \frac{1}{6D_r} \quad \text{Eq. 2.10}$$

where τ_c is the timescale of tumbling, D_r is the rotational diffusion constant, k is Boltzmann's constant, T is temperature, η is the viscosity of the solvent, and a is the radius hydration.⁴¹ The tumbling of a monomeric protein in water can be estimated as 0.6 times the molecular weight in kDa, since it is roughly spherical. The CEACAM monomer with the AMS fluorophore attached has a molecular weight of 12.59 kDa, giving it a $\tau_c \approx 7.6$ ns.

As the AMS fluorophore has a fluorescence lifetime of approximately 10 ns, which allows for sufficient rotation of the CEACAM monomer to emit depolarized light after excitation. Additionally, as we are titrating Opa proteoliposomes, the difference in molecular weight between the free CEACAM and bound state is orders of magnitude apart, which will provide us with a large change in polarization (Figure 2.6).

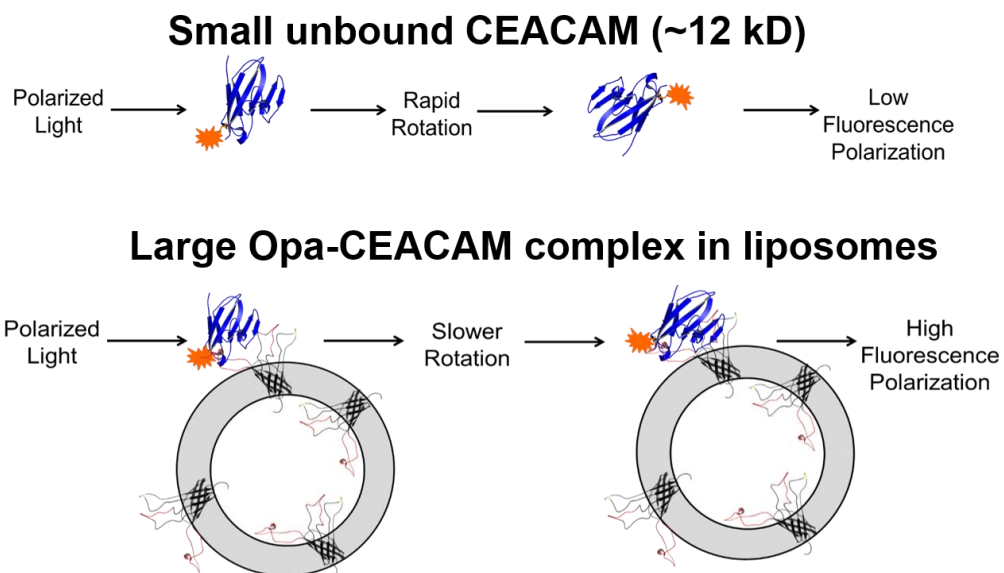


Figure 2.6. Schematic of FP with CEACAM and Opa proteoliposomes. The polarization of emitted light from fluorescently-labeled NCCM is modulated upon binding to Opa proteoliposomes.

2.4.3 Analysis of FP data can be complicated by several factors

While FP is a relatively straight forward technique to measure the interactions of molecules, analysis of the results can be complicated by several factors. A major concern for FP is fluorescent quenching (a decrease in fluorescence intensity). The change in polarization between the bound and free state is not only dependent on the tumbling, but of the environment as well. If small labeled molecule binds in a way that the fluorophore can interact, it may change the intensity of the fluorescent light emitted. Before analyzing FP data, fluorescent quenching must be assessed by determining the fluorescence intensity of the labeled ligand in a bound and unbound state.. When quenching is detected, the data must be analyzed using the following equation:

$$\text{Corrected fraction bound} = \frac{P - P_f}{(P_b - P) \left(\frac{Q_b}{Q_f} \right) + P - P_f} \quad \text{Eq. 2.11}$$

where P is the observed polarization, P_f is the polarization of the ligand in the free state, P_b is the polarization of the ligand in the bound state, Q_f is the fluorescence intensity of the ligand in the free state, and Q_b is the fluorescence intensity of the ligand in the bound state.⁴⁰

A second major concern with FP data is depletion of the ligand. Since the concentration of the ligand is held constant at a low value, if the affinity of the interaction with the receptor is extremely tight, too much of the ligand may be bound to get an accurate measurement. The previous equations used to calculate the fraction of ligand bound to the receptor assume that only a fraction of ligand is bound, so a new equation must be utilized for tight interactions:

$$Y = \left(\frac{K_D + R_T + L_T - \sqrt{(K_D + R_T + L_T)^2 - 4R_T L_T}}{2L_T} \right) \quad \text{Eq. 2.12}$$

Where Y is the fraction bound, K_D is the ligand dissociation constant, R_T is the receptor concentration, and L_T is the total fluorescent ligand concentration.⁴² Equation 2.12 is used to analyze the Opa – CEACAM FP data that is presented in chapter 3.

2.5 Biolayer interferometry (BLI) provides fast, label free assessment of interactions

2.5.1 Underlying principles of BLI

Biolayer Interferometry (BLI) is a relatively new technique used to measure interactions of biomolecules, similar to surface plasmon resonance.⁴³ The first commercial instrument became available in 2005 from fortéBIO. The technology used in these BLI instruments is based on the principles of optical interference. Two waves of the same frequency, traveling in the same direction will interfere with each other and produce a new wave, resulting from the net effect of both individual waves. This interference is considered to be constructive when both of the propagating waves are in the same phase (Figure 2.7, red), or destructive when the two propagating waves are entirely out of phase (Figure 2.7, blue). The resulting wave will have an amplitude equal to the net sum of the propagating waves (Figure 2.7).

As explained by Dayne and Do *et al*, BLI uses a glass fiber sensor with a proprietary optical layer for measuring samples (Figure 2.8).^{43,44} White light is emitted down the sensor and is reflected back towards a detector from two interfaces: (i) the optical layer and (ii) the tip of the sensor, including anything that is bound to the surface (Figure 2.8). Each wavelength of light is analyzed in terms of interference, as there are two waves traveling back up the sensor from reflection. An interference pattern emerges from the constructive and destructive interferences of all wavelengths of light.^{43,44}

When molecules bind to the tip of the biosensor, they increase the path length of the reflected light, while the path length to the optical layer remains unchanged. This changes the interference pattern, and can be measured in real time. When more molecules bind to the sensor, it shifts the interference pattern further. Molecules dissociating from the sensor can also be measured, as dissociation shifts the interference pattern.^{43,44} A schematic summary of the principles behind BLI can be found in Figure 2.9.

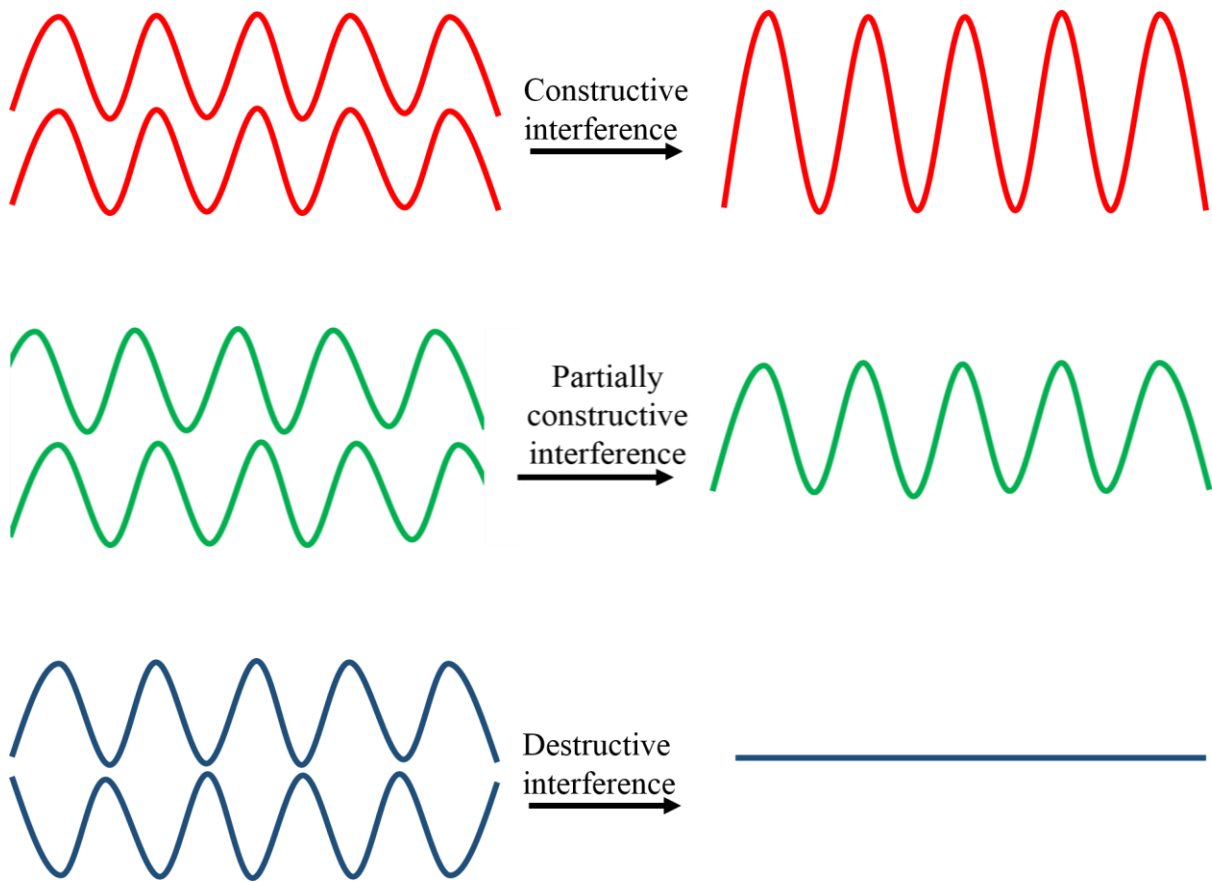


Figure 2.7. Interference patterns of waves. Constructive interference occurs when two waves are perfectly in phase, and the resulting wave is double the amplitude (top, red). When two waves are slightly out of phase, the resulting wave amplitude is equal to the net sum of the propagating waves, termed partial constructive interference (middle, green). Destructive interference results from two waves that are perfectly out of phase, resulting in a new wave with an amplitude of zero (bottom, blue).

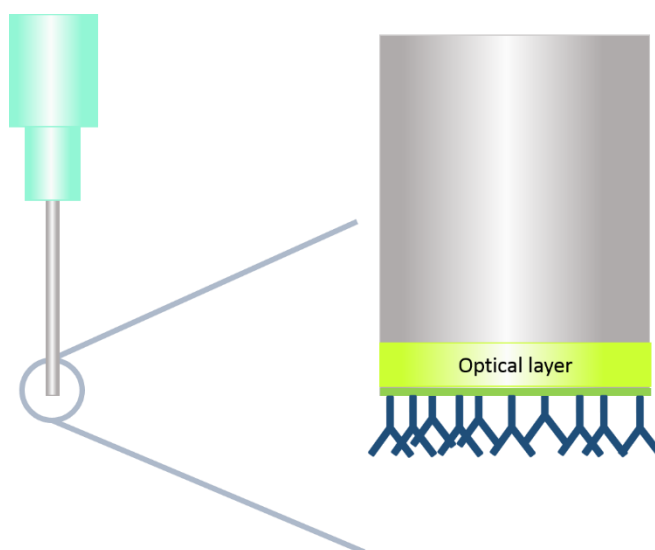


Figure 2.8. Representative surface of a biosensor. The glass fiber contains a layer with specific optical properties which allows light to pass through, as well as reflect. Molecules (typically antibodies) are immobilized on the tip of the biosensor to bind to ligands of interest.

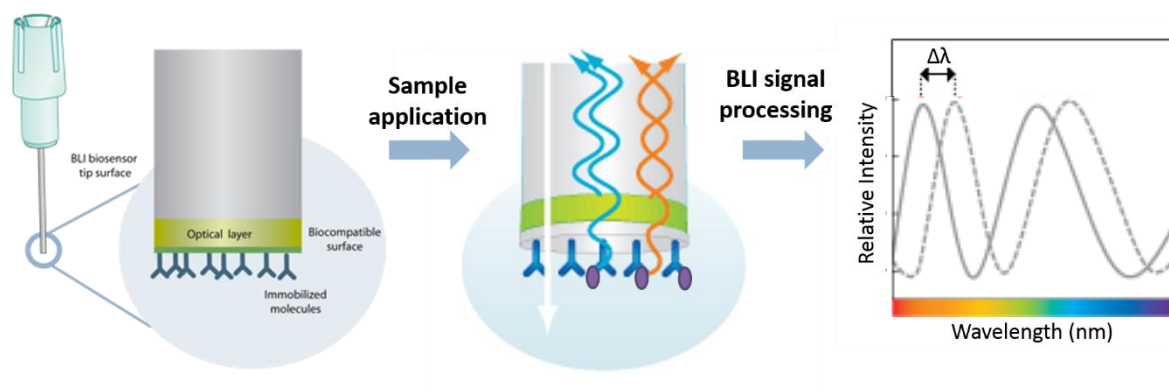


Figure 2.9. Schematic of biolayer interferometry. Adapted from Dayne, D.⁴⁴ White light is sent down the biosensor and reflected back from a reference point (the optical layer) as well as from the interface between molecules bound to the sensor and the solution. The interference of each wavelength is recorded, and shifts as more molecules bind to the surface of the sensor ($\Delta\lambda$).

2.5.2 Assessing *Opa* – CEACAM interactions using BLI

While a purported benefit of BLI is that the sample can be label free, one molecule must be immobilized on the biosensor. Therefore, one of the binding partners does require a tag. A wide selection of biosensors are available commercially from fortéBIO, including those sensitive to His tags (coated with Anti-His antibodies or Ni-NTA), GST tags (coated with Anti-GST antibodies), or Biotinylated proteins (coated with streptavidin). To measure the *Opa* – CEACAM interaction, we are utilizing a GST tagged CEACAM construct.

A summary of the BLI experiment is depicted in Figure 2.10. Before the GST-CEACAM can be immobilized on the sensor, the sensor must first be initialized with buffer, and a baseline is recorded. Next, the sensor is dipped into a well containing the GST-CEACAM protein. The GST binds to the antibodies immobilized on the tip of the sensor, anchoring the CEACAM in place. Before applying the binding partner, the sensors must be washed with buffer to allow any GST-CEACAM not tightly attached to dissociate, so unbound receptors will not interfere with the *Opa* – CEACAM interactions. After this short wash step, the sensors are dipped into a well containing *Opa* proteoliposomes, which interact with the immobilized CEACAMs. An association curve is recorded as the shift in wavelength versus time, and can be fit using the following association equation:

$$Y = A(1 - e^{-k_{obs} * t}) \quad \text{Eq. 2.13}$$

Where Y is the nm shift in wavelength (corresponds to fraction bound), A is the asymptote, k_{obs} is the observed association constant, and t is time from the start of association.⁴⁵

The final step is to dip the sensor, now containing the *Opa* proteoliposomes bound to the CEACAM, into buffer and allow the *Opa* to dissociate. A dissociation curve is recorded, from which k_{off} (1/s) is determined using the following equation:

$$Y = (Y_0 - A) * e^{-k_{off} * t} + A \quad \text{Eq. 2.14}$$

Where Y_0 is the nm shift at the start of dissociation, A is the asymptote (assumed to be zero), k_{off} is the dissociation constant, and t is time from the start of dissociation.⁴⁵

Importantly, it is noted that k_{obs} is not equivalent with k_{on} . Dissociation is occurring during the association phase, which must be taken into account. Therefore, k_{obs} is a constant which reflects the time it takes for the association to reach equilibrium. From the values of k_{off} and k_{obs} , k_{on} (1/M*s) can be calculated using the following equation:

$$k_{on} = \frac{k_{obs} - k_{off}}{[Opa]} \quad \text{Eq. 2.15}$$

The dissociation constant (K_D) is the ratio of k_{off}/k_{on} , which can then be calculated, and will provide insights towards Opa protein – CEACAM receptor interactions.

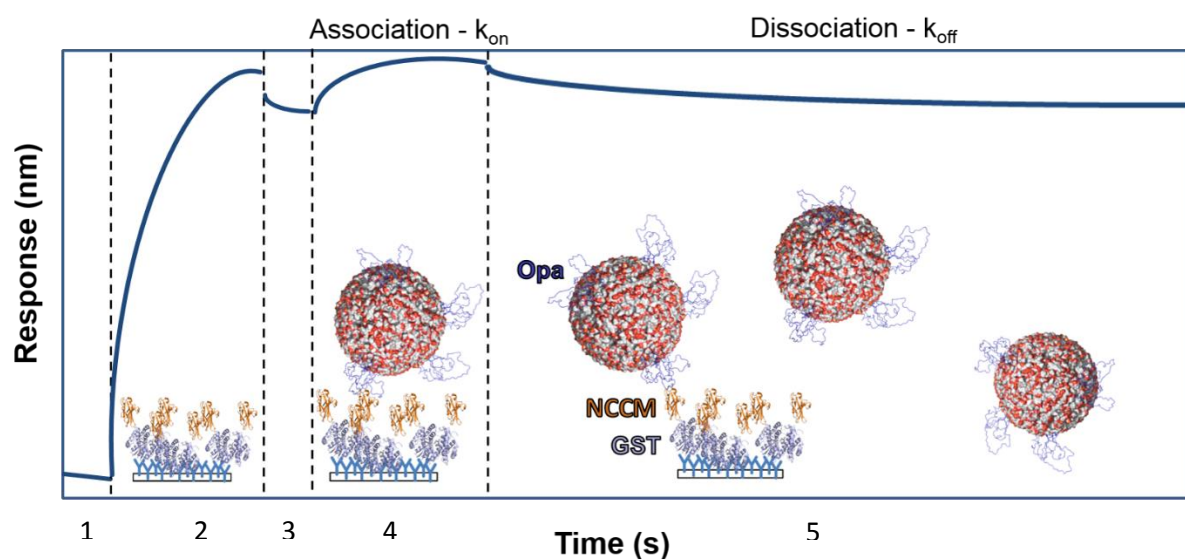


Figure 2.10. Theoretical data from a BLI experiment with GST-NCCM and Opa proteoliposomes. The first panel (1) shows a baseline of the biosensor in the buffer that will be used for the binding experiment. The second panel (2) shows the binding of GST-NCCM to the GST antibodies on the surface of the sensor. Next (3), the sensor is dipped into buffer to allow any dissociation of the GST-NCCM not tightly associated with the antibodies. The fourth panel (4) shows the binding of the Opa proteoliposomes to the NCCM attached to the tip of the sensor via the GST, which generates an association curve. Lastly (5), the sensor is again placed into buffer to allow the Opa proteoliposomes to dissociate from the NCCM, generating a dissociation curve.

2.6 References

1. Patching, S.G. NMR structures of polytopic integral membrane proteins. *Molec Membr Biol* **28**, 370-397 (2011).
2. Overington, J.P., Al-Lazikani, B. & Hopkins, A.L. How many drug targets are there? *Nat Rev Drug Discov* **5**, 993-996 (2006).
3. Liu, J. & Rost, B. Comparing function and structure between entire proteomes. *Protein Sci* **10**, 1970-1979 (2001).
4. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodospseudomonas viridis* at 3[angst] resolution. *Nature* **318**, 618-624 (1985).
5. Luckey, M. *Membrane Structural Biology with Biochemical and Biophysical Foundations*, (Cambridge University Press, 2008).
6. White, S.H. The progress of membrane protein structure determination. *Protein Sci* **13**, 1948-1949 (2004).
7. White, S. Membrane Proteins of Known 3D Structure. (2016).
8. Surrey, T. & Jähnig, F. Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Proc Natl Acad Sci U S A* **89**, 7457-7461 (1992).
9. Surrey, T. & Jähnig, F. Kinetics of Folding and Membrane Insertion of a β -Barrel Membrane Protein. *J Biol Chem* **270**, 28199-28203 (1995).
10. Pocanschi, C.L., Patel, G.J., Marsh, D. & Kleinschmidt, J.H. Curvature Elasticity and Refolding of OmpA in Large Unilamellar Vesicles. *Biophys J* **91**, L75-L77 (2006).
11. Patel, G.J., Behrens-Kneip, S., Holst, O. & Kleinschmidt, J.H. The Periplasmic Chaperone Skp Facilitates Targeting, Insertion, and Folding of OmpA into Lipid

- Membranes with a Negative Membrane Surface Potential. *Biochemistry* **48**, 10235-10245 (2009).
12. Surrey, T., Schmid, A. & Jähnig, F. Folding and Membrane Insertion of the Trimeric β -Barrel Protein OmpF. *Biochemistry* **35**, 2283-2288 (1996).
 13. Mahalakshmi, R., Franzin, C.M., Choi, J. & Marassi, F.M. NMR structural studies of the bacterial outer membrane protein OmpX in oriented lipid bilayer membranes. *Biochim Biophys Acta* **1768**, 3216-3224 (2007).
 14. Shanmugavadivu, B., Apell, H.-J., Meins, T., Zeth, K. & Kleinschmidt, J.H. Correct Folding of the β -Barrel of the Human Membrane Protein VDAC Requires a Lipid Bilayer. *J Mol Biol* **368**, 66-78 (2007).
 15. Burgess, N.K., Dao, T.P., Stanley, A.M. & Fleming, K.G. β -Barrel Proteins That Reside in the *Escherichia coli* Outer Membrane in Vivo Demonstrate Varied Folding Behavior in Vitro. *J Biol Chem* **283**, 26748-26758 (2008).
 16. Dewald, A.H., Hodges, J. C., Columbus, L. Physical determinants of beta-barrel membrane protein folding in lipid vesicles. *Biophys J* **100**, 2131-2140 (2011).
 17. Dewald, A.H. Folding and biophysical characterization of Neisserial outer membrane Opacity-associated (Opa) proteins in lipid vesicles. University of Virginia (2012).
 18. Petrache, H.I., Dodd, S.W. & Brown, M.F. Area per Lipid and Acyl Length Distributions in Fluid Phosphatidylcholines Determined by ^2H NMR Spectroscopy. *Biophys J* **79**, 3172-3192 (2000).
 19. Small, D.M. *The Physical Chemistry of Lipids: From Alkanes to Phospholipids (Handbook of Lipid Research)*, (Springer, 1986).

20. Swairjo, M.A., Seaton, B.A. & Roberts, M.F. Effect of vesicle composition and curvature on the dissociation of phosphatidic acid in small unilamellar vesicles - a ³¹P-NMR study. *Biochim Biophys Acta* **1191**, 354-361 (1994).
21. Cevc, G. & Richardsen, H. Lipid vesicles and membrane fusion. *Adv Drug Deliv Rev* **38**, 207-232 (1999).
22. Bittman, R. A review of the kinetics of cholesterol movement between donor and acceptor bilayer membranes. in *Cholesterol in membrane models* (ed. Finegold, L.X.) (CRC Press, 1992).
23. Du, H., Chandaroy, P. & Hui, S.W. Grafted poly-(ethylene glycol) on lipid surfaces inhibits protein adsorption and cell adhesion. *BBA Biomembranes* **1326**, 236-248 (1997).
24. Price, M.E., Cornelius, R.M. & Brash, J.L. Protein adsorption to polyethylene glycol modified liposomes from fibrinogen solution and from plasma. *BBA Biomembranes* **1512**, 191-205 (2001).
25. Green, A.E. & Rose, P.G. Pegylated liposomal doxorubicin in ovarian cancer. *Int J Nanomedicine* **1**, 229-239 (2006).
26. Criss, A.K. & Seifert, H.S. A bacterial siren song: intimate interactions between *Neisseria* and neutrophils. *Nat Rev Micro* **10**, 178-190 (2012).
27. Fox, D.A. et al. Structure of the Neisserial Outer Membrane Protein Opa60: Loop Flexibility Essential to Receptor Recognition and Bacterial Engulfment. *JACS* **136**, 9938-9946 (2014).
28. Woodle, M.C. & Lasic, D.D. Sterically stabilized liposomes. *Biochim Biophys Acta* **1113**, 171-99 (1992).

29. Blume, G. & Cevc, G. Molecular mechanism of the lipid vesicle longevity *in vivo*. *Biochim Biophys Acta* **1146**, 157-68 (1993).
30. Vert, M. & Domurado, D. Poly(ethylene glycol): protein-repulsive or albumin-compatible? *J Biomater Sci Polym Ed* **11**, 1307-17 (2000).
31. Immordino, M.L., Dosio, F. & Cattell, L. Stealth liposomes: review of the basic science, rationale, and clinical applications, existing and potential. *Int J Nanomedicine* **1**, 297-315 (2006).
32. Krown, S.E., Northfelt, D.W., Osoba, D. & Stewart, J.S. Use of liposomal anthracyclines in Kaposi's sarcoma. *Semin Oncol* **31**, 36-52 (2004).
33. Needham, D., McIntosh, T.J. & Lasic, D.D. Repulsive interactions and mechanical stability of polymer-grafted lipid membranes. *Biochim Biophys Acta* **1108**, 40-8 (1992).
34. de Gennes, P.G. Conformations of Polymers Attached to an Interface. *Macromolecules* **13**, 1069-1075 (1980).
35. Hagan, C.L., Silhavy, T.J. & Kahne, D. β -Barrel Membrane Protein Assembly by the Bam Complex. *Annu Rev Biochem* **80**, 189-210 (2011).
36. Fox, D.A. & Columbus, L. Solution NMR resonance assignment strategies for β -barrel membrane proteins. *Protein Sci* **22**, 1133-1140 (2013).
37. Woodbury, C.P. *Introduction to Macromolecular Binding Equilibria*, (Taylor & Francis, 2008).
38. Lakowicz, J.R. *Principles of Fluorescence Spectroscopy*, (Plenum Publishers, 1999).
39. Valeur, B., Berberan-Santos, M. *Molecular fluorescence: principles and applications*, (Wiley, 2012).

40. *Fluorescence polarization technical resource guide*, (Invitrogen Corporation).
41. Lilly, A.A., Crane, J.M. & Randall, L.L. Export chaperone SecB uses one surface of interaction for diverse unfolded polypeptide ligands. *Protein Sci* **18**, 1860-8 (2009).
42. Veiksina, S., Kopanchuk, S. & Rinken, A. Fluorescence anisotropy assay for pharmacological characterization of ligand binding dynamics to melanocortin 4 receptors. *Anal Biochem* **402**, 32-39 (2010).
43. Do, T. et al. A rapid method for determining dynamic binding capacity of resins for the purification of proteins. *Protein Express Purif* **60**, 147-150 (2008).
44. Dayne, D. BioLayer Interferometry (BLI) - How Does it Work? *ForteBIO Interactions* **5**, 5 (2012).
45. Wilson, J.L., Scott, I.M. & McMurry, J.L. Optical biosensing: Kinetics of protein A-IGG binding using biolayer interferometry. *BAMBED* **38**, 400-407 (2010).

3. Reconstituted Opa proteoliposomes are an effective tool for studying Opa protein interactions *in vitro*

3.1 Overview

Membrane protein research is often hindered by the need to find an appropriate *in vitro* environment that allows the protein to remain soluble and functional. Even though a membrane protein may be soluble in a wide variety of membrane mimics, it does not guarantee that the protein will be functional. Additional investigations must be undertaken to ensure that the *in vitro* recombinant membrane protein maintains functionality similar to that of the *in vivo* protein.

To compare the function of Opa proteins *in vivo* and *in vitro*, binding was assessed qualitatively and quantitatively between a selection of Opa proteins (Opa₆₀, OpaD, and Opa₅₀), either expressed in Gc or reconstituted into liposomes, and CEACAM receptors (CEACAM1 and CEACAM3). Pull-down assays, coupled with immunoblotting, were utilized to qualitatively determine that recombinant Opa proteins in liposomes retain the ability to interact with receptors, as compared to Opa proteins expressed in Gc. This suggests a conservation of Opa protein structure outside of Gc, and establishes a platform for future *in vitro* investigations. In order to quantify the recombinant Opa – CEACAM interaction, fluorescence polarization was used, and a tight (nM) affinity was found for all Opa_{CEA} protein – CEACAM combinations investigated. We hypothesize that such a tight affinity between Opa proteins and CEACAM receptors is necessary for effective competition with all of the other interactions that CEACAMs are capable of forming *in vivo*.

3.2 Introduction

3.2.1 *Neisserial attachment to host cells*

The genus *Neisseria* contains both pathogenic and commensal species which colonize human epithelia.¹ The commensal species live on the mucosal surface of the nasopharynx, but seldom cause the host any discomfort or illness.¹ However, the pathogenic *Neisseria gonorrhoeae* (Gc) and *Neisseria meningitidis* (Nm) can cause inflammation by their colonization, as well as induce engulfment of the bacteria.¹ While the pathogenic and commensal species of *Neisseria* are very similar genetically, the pathogens Gc and Nm express additional factors that allow efficient colonization of the mucosa, evasion of the immune system, and effective transmission.¹

The relationship between the pathogenic *Neisseria* and the immune system is complex. *Neisseria* employ a number of mechanisms that enable the bacteria to evade the immune system.¹ The phase variation and sequence variability of Opa proteins are one such mechanism. Neutrophils (polymorphonuclear leukocytes, PMNs) are the immune system's first responders to injury or infection, and stream to the site of Gc or Nm infection.¹ Neutrophils are able to kill bacteria extracellularly, as well as by phagocytosis.² Neutrophils possess a collection of proteins that are able to mediate bacterial engulfment, including CEACAMs 1 and 3.³

3.2.2 *CEACAM as an adhesion molecule in cells*

The ability of cells to interact not only with each other, but with their environment as well is critical for maintaining proper function. Aberrant cell adhesion is indicated as playing a pivotal role in the development and progression of cancer.⁴ The immunoglobulin superfamily of cell adhesion molecules (Ig-CAMs) have been one of several classes of molecules

that are highly researched as being involved in the spread of certain cancers.⁴ The carcino-embryonic adhesion (CEA) family of proteins are members of the Ig-CAM super-family, and have been implicated in both homotypic and heterotypic interactions.⁵⁻¹⁰ Many bacterial pathogens express proteins that interact with NCCMs, especially CEACAM1, 3, 5, and 6.¹¹⁻¹⁶ Additionally, NCCMs are known to interact specifically with fimbrial structures such as Dr adhesins.⁸ Interactions with NCCMs may occur through carbohydrate moieties on the NCCM, as is the case for CEACAM1 binding to enterobacteria, such as *Escherichia coli* and *Salmonella* strains.³ CEACAMs also interact through a non-glycosylated region of NCCM, as is observed for *Haemophilus influenzae*, *Moraxella catarrhalis*, and the pathogenic *Neisseria* species.¹⁷

3.2.3 CEACAMs as a receptor for *Neisseria* Opa proteins

Interactions between CEACAM and the pathogenic *Neisseria* not only allow the bacteria to adhere to and colonize human cells, but can also trigger engulfment of the bacteria. These interactions have been studied in epithelial cells and PMNs. Epithelial cells express the pathogen binding CEACAMs 1, 5, and 6, as well as the non-pathogen binding CEACAM7.^{18,19} Primary human neutrophils express CEACAMs 1, 3 and 6, as well as the non-pathogen binding CEACAMs 4 and 8.²⁰ Of particular interest are CEACAMs 1 and 3; both contain cytoplasmic domains involved in signaling that can lead to the internalization of the bacterium, but often trigger opposing cell responses.

CEACAM1 contains two immunoreceptor tyrosine-based inhibition motifs (ITIMs) which upon activation triggers the recruitment of the phosphatase SHP-1 to suppress phosphotyrosine-based signaling cascades.²¹⁻²³ CEACAM3 contains an immunoreceptor

tyrosine-based activation motif (ITAM), which recruits kinases (such as Syk, the Src family kinases) upon activation to propagate pro-inflammatory signaling cascades.²⁴⁻²⁷ CEACAM3, which is expressed exclusively on human neutrophils and other granulocytes, is thought to have evolved as innate immune protection, as it has no known endogenous ligand, but interacts specifically with proteins expressed on the surface of human-specific bacterial pathogens such as *Neisseria*. These CEACAM3 interactions mediate uptake of the pathogen, which induces killing of the bacterium via oxidative burst as well as toxic granule release.^{3,28}

CEACAM receptors mediate Gc and Nm engulfment by binding to Neisserial opacity-associated (Opa) proteins. The majority of Opa proteins interact with CEACAMs, however, Opa_{CEA} proteins vary in their specificity of interactions with CEACAMs.¹¹ All Opa_{CEA} proteins interact with the non-glycosylated face of the IgV-like domain of CEACAMs.²⁹ Ten CEACAM residues were identified which mediate binding to Opa proteins; of these important residues, only Tyr68 and Ile125 (residue numbers of CEACAM1, UniProt ID P13688-1) interact with all studied Opa variants, and are highly conserved on all CEACAMs.³⁰ Of the other eight CEACAM residues involved in binding to Opa proteins, six are conserved between CEACAM1 and 3 (structure, Figure 3.1, sequences Figure 3.2).

While the amino acid residues of CEACAM involved in interactions with Opa proteins are known, it is unclear which residues on Opa are involved in these interactions. We seek to elucidate the molecular determinants of Opa – CEACAM interactions. To do this, we need an *in vitro* system that mimics the *in vivo* function of Opa proteins. Towards this end, in collaboration with the laboratory of Professor Alison Criss (UVA Department of Microbiology, Immunology, and Cancer biology), we have investigated receptor specificity of several Opa proteins (Opa₆₀, an Opa_{CEA} from Gc strain MS11, OpaD, an Opa_{CEA} from Gc strain

FA1090, and Opa₅₀, an Opa_{HS} from Gc strain MS11, sequences in Figure 3.3) towards NCCM1 and NCCM3.

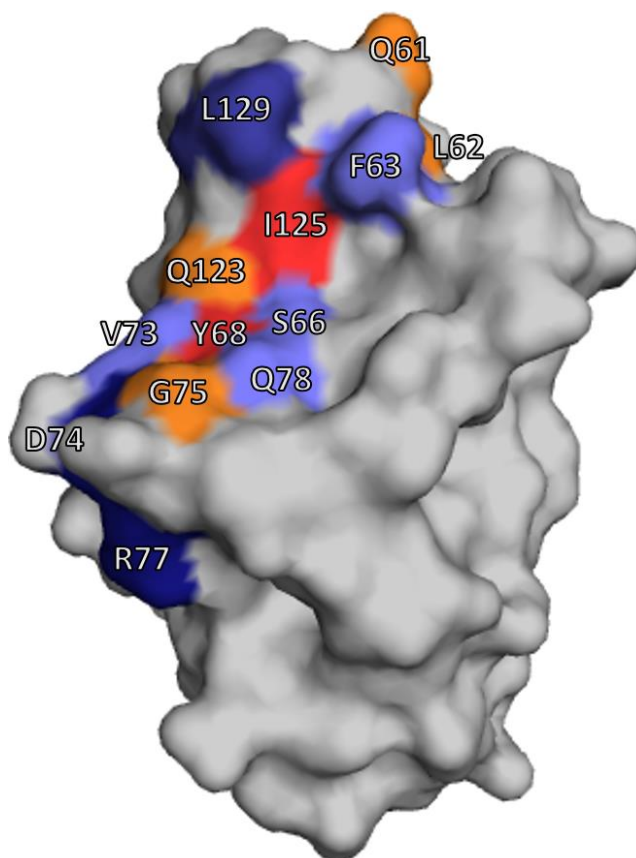


Figure 3.1. Surface representation of NCCM1. Residues shown in red are involved in all interactions with Opa_{CEA} proteins. Residues in orange only participate in binding to specific Opa variants. Residues shown in dark blue are involved in NCCM homotypic and heterotypic interactions. Residues highlighted in light blue are also involved in homotypic and heterotypic interactions, and interactions with specific Opa proteins. Residue numbering corresponds to CEACAM1 sequence, UniProt ID P13688-1.

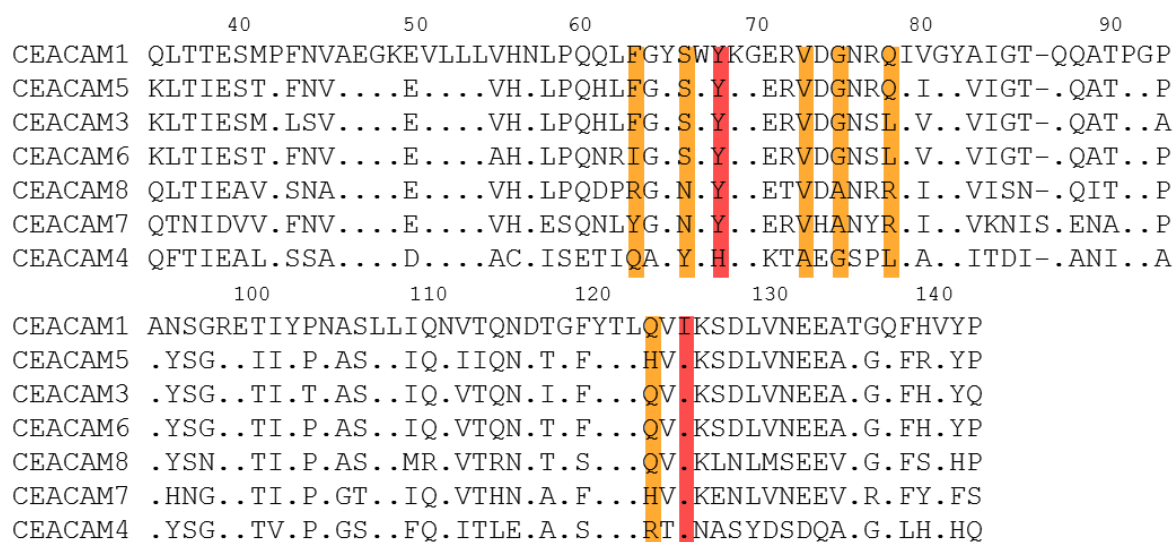


Figure 3.2. Sequence alignment of human NCCMs 1, 3, 4, 5, 6, 7, and 8. Essential residues for binding all Opa proteins are highlighted with red boxes and residues that are important to some Opa protein interactions with orange boxes. Residue numbering corresponds to CEACAM1 sequence, UniProt ID P13688-1

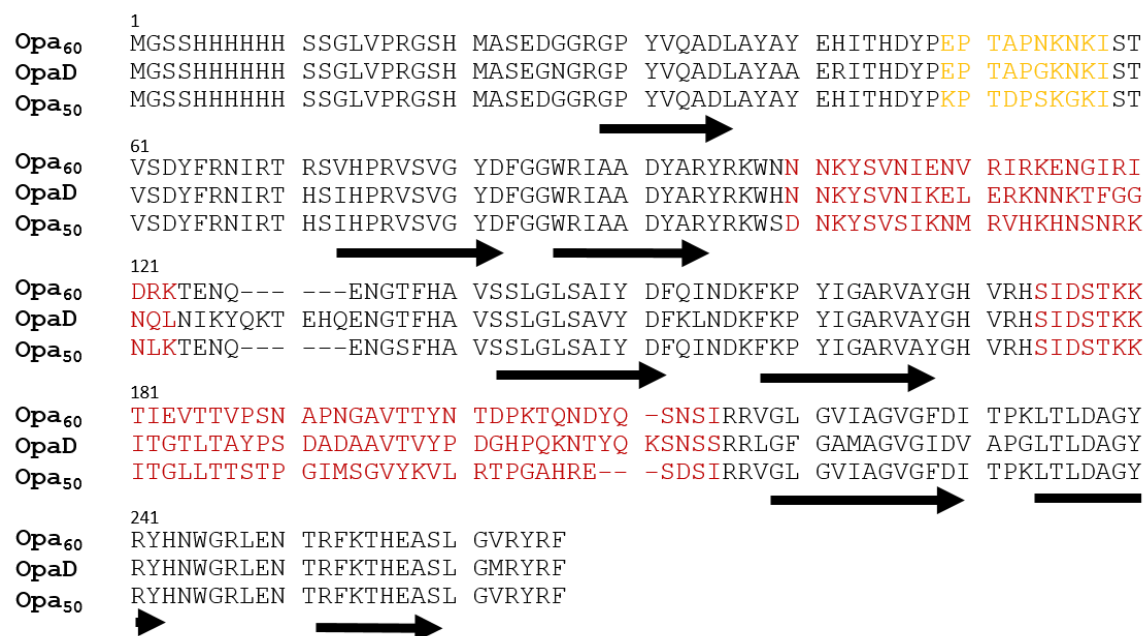


Figure 3.3. Sequence alignment of Opa₆₀ from Gc strain MS11, OpaD from Gc strain FA1090, and Opa₅₀ from Gc strain MS11. Semivariation regions are highlighted in yellow and hypervariation regions are highlighted in red. Black arrows indicate β -strands which form the membrane spanning barrel.

3.3 Results and discussion

3.3.1 Recombinant NCCM1 and 3 interact with Opa₆₀ and OpaD+ Gc

Gc that are phase ON for expression of Opa₆₀ have previously been reported to interact with both CEACAM1 and CEACAM3, both in cells transfected to express a specific CEACAM, as well as recombinant CEACAM fusion products.^{31,32} In collaboration with Louise Ball from the Criss laboratory, we used Gc constitutively expressing only Opa₆₀, OpaD, or Opa₅₀ (a heparan sulfate proteoglycan binding Opa), along with Opa-deficient (Opa⁻) bacteria, to assess the ability of Gc to associate with recombinant NCCM1 and 3.³³ In the case of NCCM3, precipitation was observed at concentrations greater than ~10 nM and the GST-NCCM3 fusion was used for all pull-down assays. Purified NCCM1 or GST-NCCM3 (referred from here on as NCCM3) was incubated with each strain of Gc, and the cells with bound CEACAM were collected using centrifugation. The fraction of NCCM bound (pellet) and unbound (supernatant) to Gc were determined using SDS-PAGE and immunoblotting.

As expected, supernatants from all Gc-NCCM combinations reacted with a pan-CEACAM antibody, and all pellets containing Opa⁺ bacteria reacted with a pan-Opa antibody (Figure 3.4). NCCM1 and NCCM3 associated with both Opa₆₀ and OpaD Gc, as assessed by the presence of NCCM in the bacterial pellet. In comparison, Opa₅₀ and Opa⁻ Gc showed negligible association with both NCCMs. To confirm that the GST fusion did not interfere with binding, NCCM1 and GST-NCCM1 binding to Opa expressing Gc were compared and found to be similar (Figure 3.5). These findings validate the pull-down assay for the specificity and selectivity of Opa_{CEA} – NCCM interactions, based on previous reports with MS11 Opa₅₀ and Opa₆₀,^{31,32,34} and furthermore show that OpaD binds CEACAM3 to a similar extent as CEACAM1.

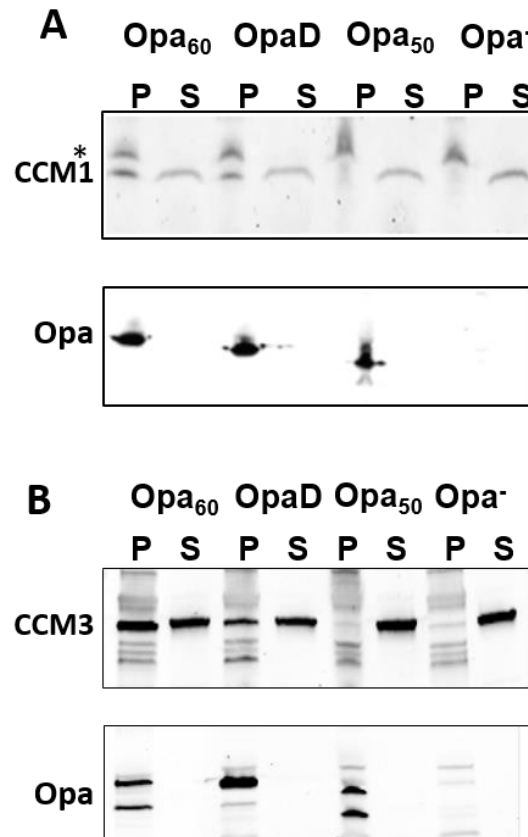


Figure 3.4. CEACAMs interact specifically with Opa_{CEA} proteins expressed in Gc. Bacteria expressing OpaD, Opa₅₀, or Opa₆₀, or Opa⁻ Gc, were incubated with NCCM1 (A) or GST-NCCM3 (B), and samples were centrifuged for pellet (P) and supernatant (S) immunoblot assessment. For Opa immunoblots, the two bands correspond to folded (lower band) and unfolded (upper band) protein. *Higher molecular weight bands in the Gc pellet samples of the CEACAM blots indicate nonspecific CEACAM antibody reactivity with antigens on the surface of Gc.

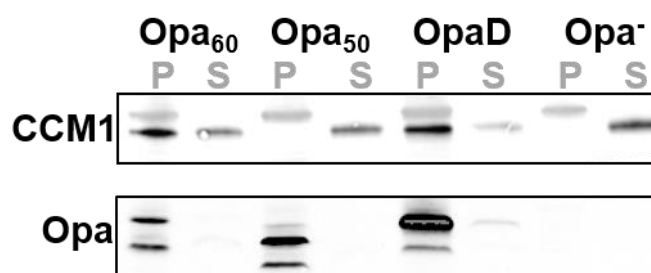


Figure 3.5. Immunoblot of pull-down assay using Opa expressing Gc and GST-NCCM1.

Opa₆₀ and OpaD Gc interacted with GST-NCCM1 similarly to NCCM1 (compare to Figure 2) indicating the GST tag does not interfere with NCCM binding to Opa proteins.

3.3.2 Recombinant *Opa*₆₀ and *OpaD* reconstituted in liposomes interact with recombinant NCCM1 and 3

To determine if Opa proteins retain their receptor specificity *in vitro*, pull-down assays were performed with Opa proteins reconstituted into small unilamellar liposomes (less than 100 nm)³⁵. Bound NCCM was evaluated by comparing the supernatant and pellet after ultracentrifugation of the Opa proteoliposome – NCCM mixtures, using SDS-PAGE and immunoblotting. In addition to Opa₅₀, trypsin treated Opa₆₀ (Opa_{Trp}) was used as a negative control instead of liposomes without protein which are too buoyant to be pelleted by centrifugation. The supernatants from each proteoliposome – NCCM mixture reacted with the CEACAM antibody, while all of the pellets reacted with the Opa antibody (except for Opa_{Trp}, which likely does not contain the Opa antibody epitope) (Figure 3.6). Pellets for Opa₆₀ and OpaD proteoliposome contained both NCCM1 and 3, indicating that these Opa proteins retain their ability to interact with both NCCM1 and 3 when recombinantly expressed and reconstituted into liposomes.

For NCCM1, weaker bands are observed for Opa₅₀ and Opa_{Trp} compared to Opa₆₀ and OpaD indicating that there may be some non-specific interaction of NCCM1 with the liposomes since the band was also observed in the Opa_{Trp} liposomes. In contrast, the NCCM3 band observed in the pellet with Opa₅₀ contained less GST-NCCM3 than both Opa_{CEA} proteins, but more than with Opa_{Trp} indicating that Opa₅₀ demonstrates different selectivity *in vitro* compared to that observed in Gc. Together, these results suggest that the Opa protein structure and function is maintained in the liposome environment, but care needs to be taken in terms of the biological implications of the interactions investigated *in vitro* as selectivity is not conserved.

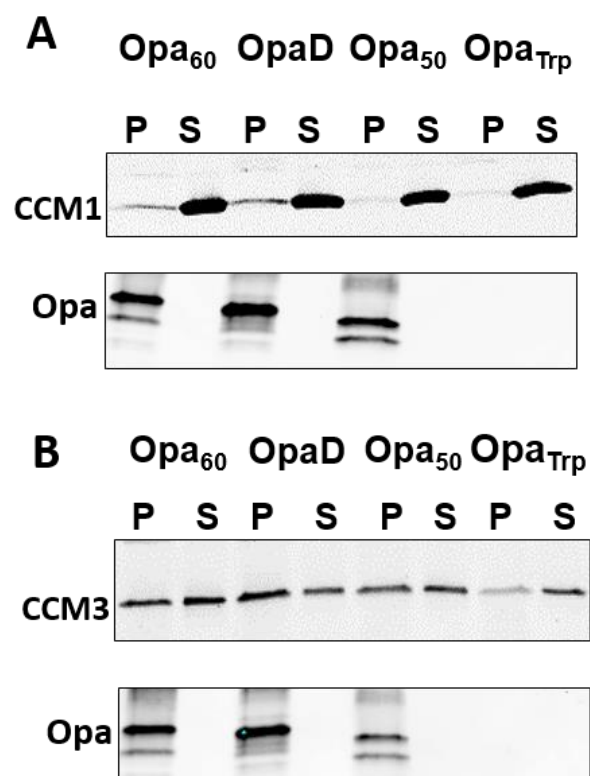


Figure 3.6. Recombinant Opa_{CEA} proteins retain their ability to interact with CEACAM.

Opa proteins were recombinantly expressed, purified, and refolded into liposomes, and then incubated with NCCM1 (A) and GST-NCCM3 (B). Samples were then centrifuged and the pellet (P) and supernatant (S) were assessed for the presence of Opa and NCCM. For Opa immunoblots, the two bands correspond to folded (lower band) and unfolded (upper band) protein.

3.3.3 NCCM1 and 3 interact with Opa₆₀ and OpaD with nanomolar affinity

Fluorescence polarization was used to quantify the strength of the interaction between Opa proteins and NCCM1 and 3. These experiments were carried out using NCCMs without the GST tag because of the lower concentrations (5 nM) of protein used in the experiment compared to the centrifugal pull-down assays. NCCM1 H139C and NCCM3 H139C were fluorescently labeled using the AMS fluorophore (described in chapter 2, Figure 3.7). The NCCM domain has no naturally occurring cysteine residue, so a cysteine mutant was introduced on the opposite side of the Opa-binding face of NCCM (Figure 3.8).

Fluorescently labeled NCCM1 or NCCM3 were incubated with Opa₆₀ or OpaD proteoliposomes or liposomes without protein. The affinities of Opa₆₀ for NCCM1 and NCCM3 were calculated to be 1.6 ± 0.6 nM and 4.3 ± 2.8 nM, respectively, while the affinities of OpaD for NCCM1 and 3 were 2.6 ± 1.3 nM and 6.8 ± 2.2 nM, respectively (Figure 3.9). The affinity of NCCM for liposomes that did not contain any Opa protein was ≥ 800 nM (data not shown). To demonstrate the interaction measured was specific to NCCM1, the fluorescence polarization experiments were repeated for Opa₆₀ with a mixture of fluorescently labeled and unlabeled proteins and the K_d values scaled with the dilution of the fluorescently labeled protein (Figure 3.10). Thus, the interactions are high affinity and similar for the combinations of the Opa – NCCM interactions investigated.

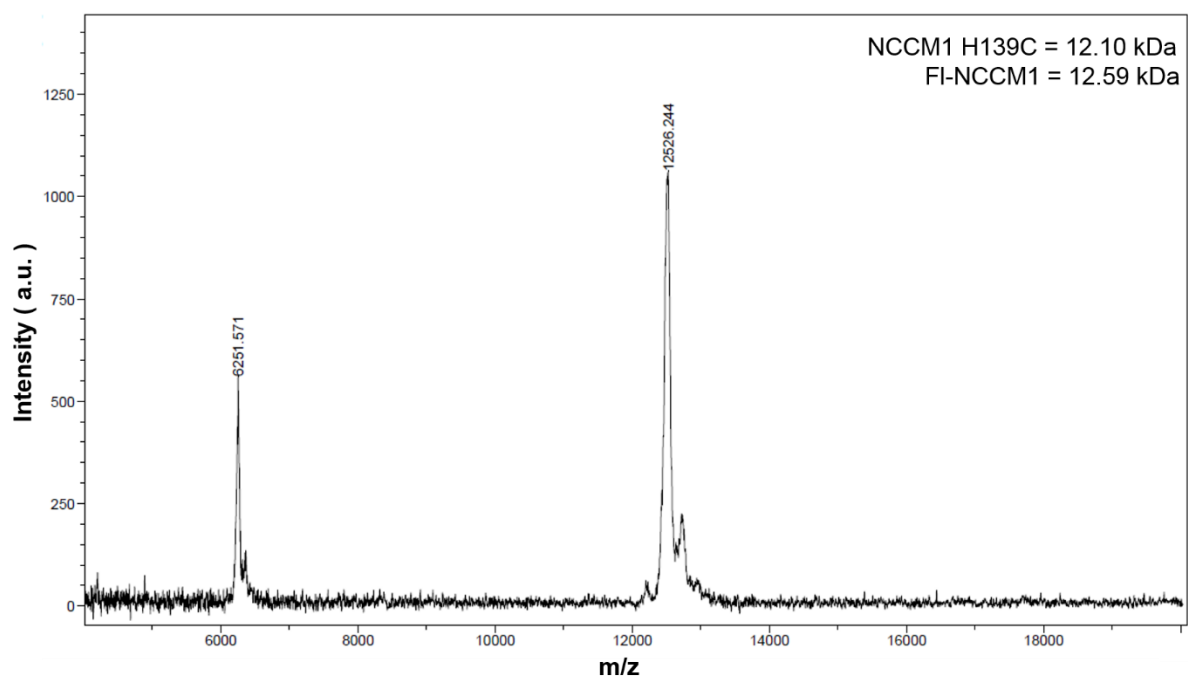
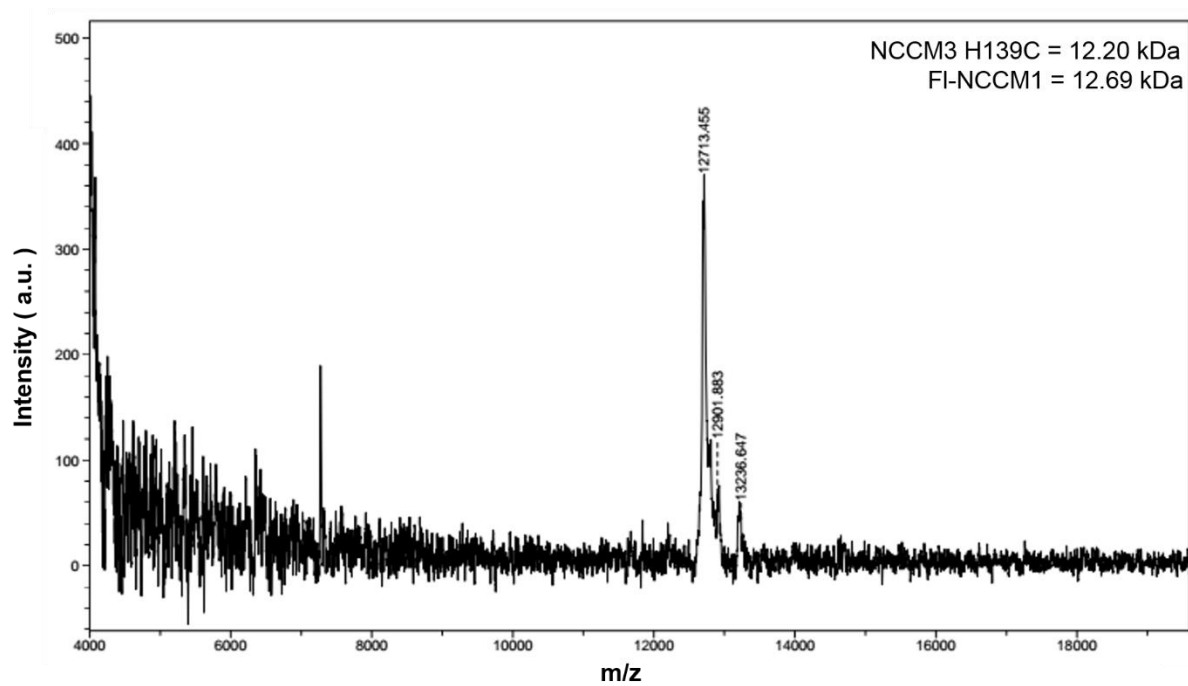
A**B**

Figure 3.7. MALDI spectra of fluorescently labeled (FI-) NCCM1 (A) and NCCM3 (B).

These spectra indicate that there is >95 % labeling efficiency of the AMS fluorophore with both NCCM1 and NCCM3.

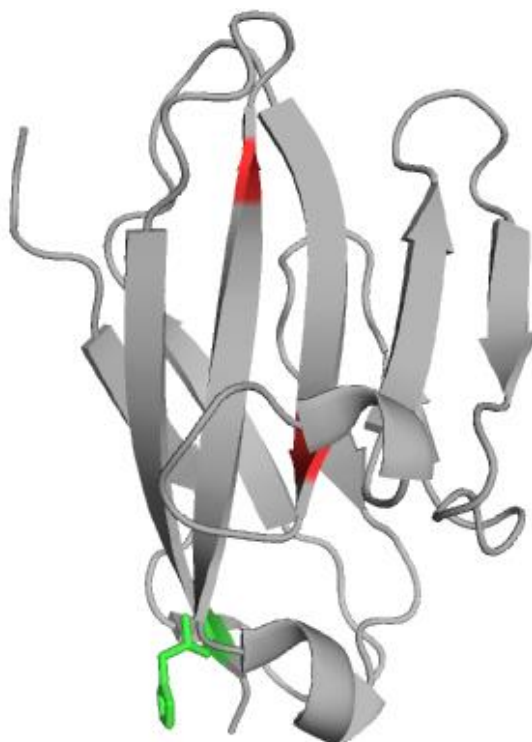


Figure 3.8. Structure of NCCM1 H139C. Ribbon diagram of NCCM1²⁹ is shown in grey, residues shown in red are Y65 and I125, which are involved in all interactions with Opa proteins (see Figure 3.1). The introduced Cys mutation (H139C) for attachment of the thiol-linked fluorophore, is on the opposite face of the NCCM1 (green).

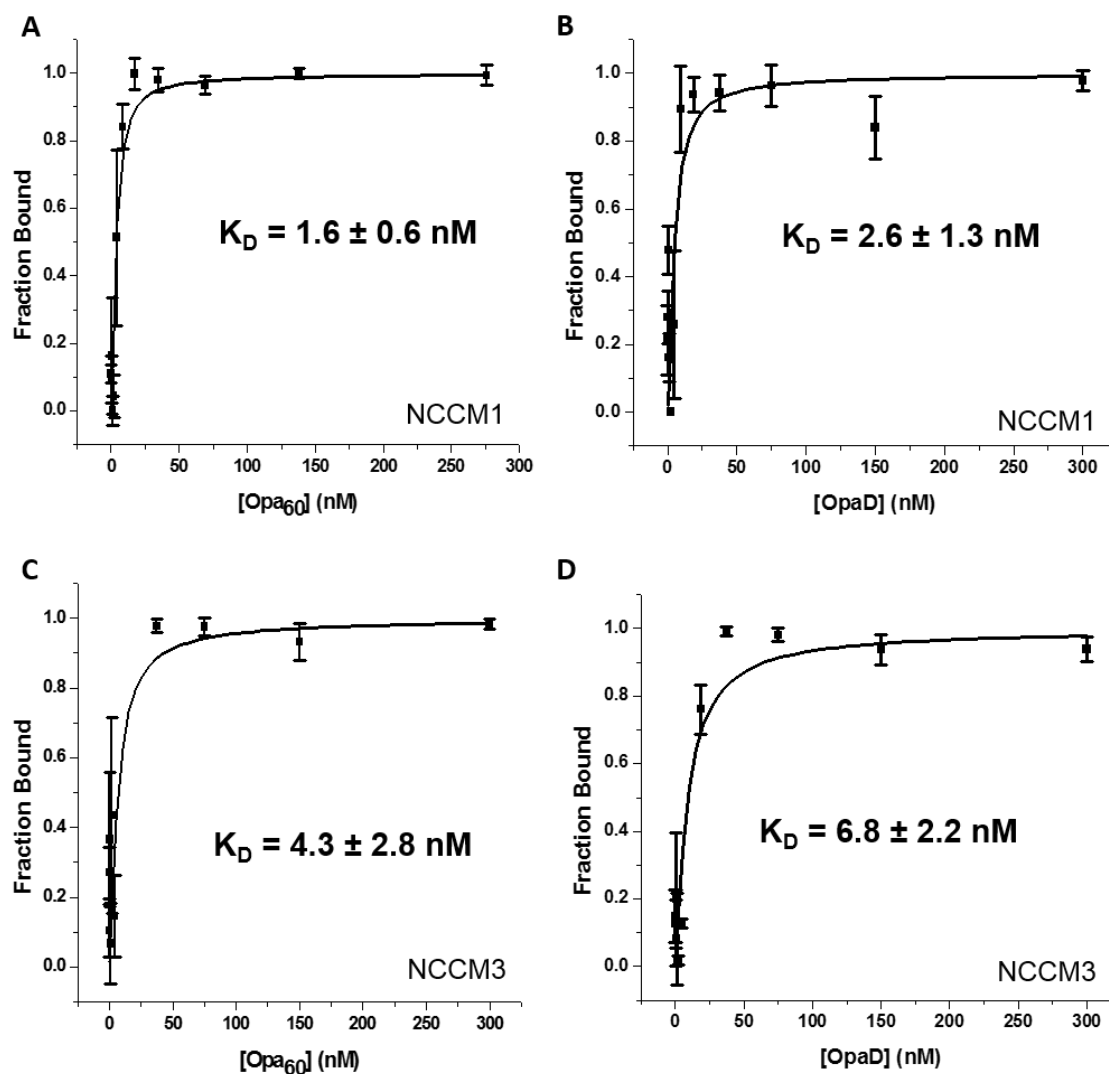


Figure 3.9. Opa_{CEA} proteins in liposomes interact with NCCM with a nanomolar affinity.

Opa₆₀ (A and C) and OpaD (B and D) were incubated with fluorescently labeled NCCM1 (A and B) or NCCM3 (C and D). Fluorescence polarization was recorded and converted into the fraction of CEACAM bound and plotted. Data was fit in OriginPro using eq. 2.12.

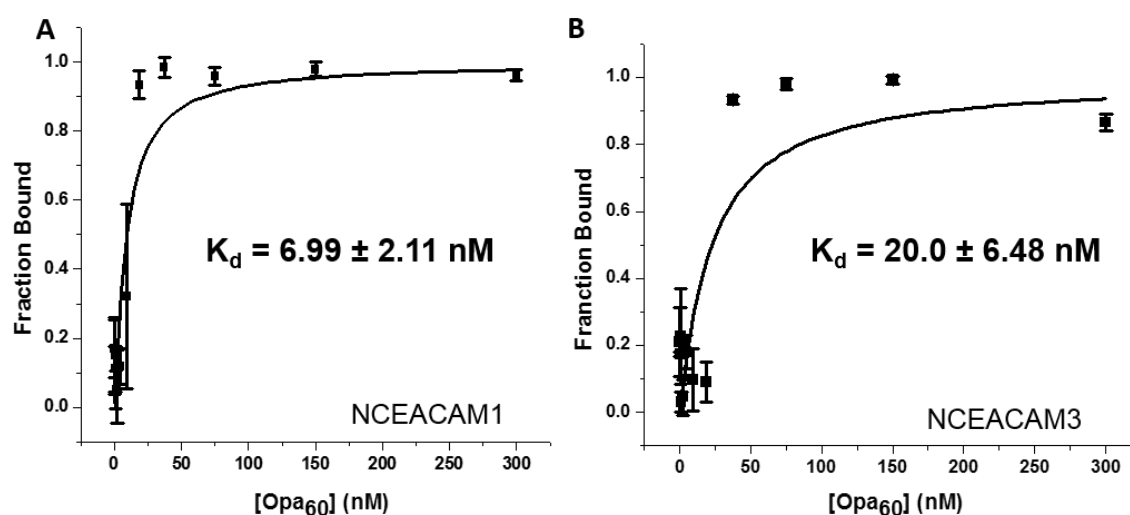


Figure 3.10. Competition FP experiments with unlabeled NCCMs 1 and 3. Opa₆₀ was incubated with 2.5 nM fluorescently labeled and 2.5 nM unlabeled NCCM1 (A) or NCCM3 (B). The determined K_d values are approximately two-fold higher than the K_d determined with 5 nM fluorescently labeled NCCM1 or 3 (Figure 3.9) indicating that the Opa protein – CEACAM interaction is specific and not due to interactions with the fluorophore.

3.3.4 Competition for CEACAM may explain high affinities

In order to effectively compete with the multiple interactions that CEACAM can be involved with, some of high affinity, the Opa – CEACAM interaction must have a tighter affinity. NCCM is directly involved in cell-cell adhesions, and many of the members of the CEACAM family can be involved in either homotypic or heterotypic interactions.^{5,6,9} Often these heterotypic interactions occur between different CEACAM variants³, but they can also involve other CEACAM domains and other molecules.^{3,8,31} Many of these interactions involve the same binding face of the IgV-like NCCM domain where Opa proteins are known to interact (Figure 3..1).^{8,9,30}

The direct hetero-interaction of NCCM5 and fibronectin was found to have a $K_d = 16$ nM.⁷ Abul-Wahid and colleagues calculated the affinity of the A₃ soluble IgC-like domain of CEACAM5 to interact with the IgV-like N-terminal domain with an affinity of 18 nM.⁷ The hetero-dimerization of NCCM6 with NCCM8 is 2 μ M.⁹ NCCM6 forms a homodimer as well, but with a much lower affinity (60 μ M).⁹ The homo-dimerization of NCCM5 has been reported with two different affinities: approximately 100 nM⁷ and 1.3 μ M⁹, both of which are tighter than other typical adhesion proteins, such as cadherins.³⁶ The affinity for the homo-dimerization of CEACAM3 is not known, however, the homo-dimerization of NCCM1 is 450 nM.⁹ Gc has many proteins that mediate adhesion to host cell membranes including adhesins and pili, which could contribute indirectly to high affinity interactions and reduce the need for direct interactions between Opa and CEACAMs to be high affinity. However, all of these surface proteins on Gc may be competing against each other to bind to their cognate receptor, which may explain why the Opa – CEACAM interaction is of such high affinity. Additionally, on the host cell side, Opa proteins need to compete with many other human proteins that

interact with CEACAMs, and potentially with other pathogens as well, which may also contribute to the high affinity of the Opa protein – CEACAM interaction.

3.4 Concluding remarks

We have demonstrated that recombinant Gc Opa proteins reconstituted into small unilamellar vesicles retain their ability to interact with CEACAM receptors, suggesting a conservation of structure and function independent of the cellular environment. Thus, we have developed an appropriate *in vitro* system to be able to further study Opa protein interactions with receptors. Furthermore, using these Opa-liposomes, we determined that two Opa_{CEA} (Opa₆₀ from Gc strain MS11 and OpaD from Gc strain FA1090) proteins have high affinity (nM) for NCCM1 and 3. While these tight affinities are not necessary for Gc attachment to host cells, we hypothesize that this tight interaction is necessary for competing with the homotypic (CEACAM – CEACAM) and heterotypic (CEACAM – other adhesion molecule) interactions of CEACAMs.

3.5 References

1. Criss, A.K. & Seifert, H.S. A bacterial siren song: intimate interactions between *Neisseria* and neutrophils. *Nat Rev Micro* **10**, 178-190 (2012).
2. Urban, C.F., Lourido, S. & Zychlinsky, A. How do microbes evade neutrophil killing? *Cell Microbiol* **8**, 1687-96 (2006).
3. Kuespert, K., Pils, S. & Hauck, C.R. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol* **18**, 565-571 (2006).
4. Okegawa, T., Pong, R-C., Li, Y., and Hsieh, J-T. The role of cell adhesion molecule in cancer progression and its application in cancer therapy. *Acta Biochim Pol* **51**, 445-457 (2004).
5. Öbrink, B. CEA adhesion molecules: multifunctional proteins with signal-regulatory properties. *Curr Opin Cell Biol* **9**, 616-626 (1997).
6. Gray-Owen, S.D. & Blumberg, R.S. CEACAM1: contact-dependent control of immunity. *Nat Rev Immunol* **6**, 433-446 (2006).
7. Abdul-Wahid, A., Huang, E.H.B., Cydzik, M., Bolewska-Pedyczak, E. & Gariépy, J. The carcinoembryonic antigen IgV-like N domain plays a critical role in the implantation of metastatic tumor cells. *Mol Oncol* **8**, 337-350 (2014).
8. Korotkova, N. et al. Binding of Dr adhesins of *Escherichia coli* to carcinoembryonic antigen triggers receptor dissociation. *Mol Microbiol* **67**, 420-434 (2008).
9. Bonsor, D.A., Günther, S., Beadenkopf, R., Beckett, D. & Sundberg, E.J. Diverse oligomeric states of CEACAM IgV domains. *Proc Natl Acad Sci U S A* **112**, 13561-13566 (2015).

10. Huang, Y.-H. et al. CEACAM1 regulates TIM-3-mediated tolerance and exhaustion. *Nature* **517**, 386-390 (2015).
11. McCaw, S.E., Liao, E.H. & Gray-Owen, S.D. Engulfment of *Neisseria gonorrhoeae*: Revealing Distinct Processes of Bacterial Entry by Individual Carcinoembryonic Antigen-Related Cellular Adhesion Molecule Family Receptors. *Infect Immun* **72**, 2742-2752 (2004).
12. Leusch, H.G., Drzeniek, Z., Markos-Pusztai, Z. & Wagener, C. Binding of *Escherichia coli* and *Salmonella* strains to members of the carcinoembryonic antigen family: differential binding inhibition by aromatic alpha-glycosides of mannose. *Infect Immun* **59**, 2051-7 (1991).
13. Virji, M. et al. Carcinoembryonic antigens are targeted by diverse strains of typable and non-typable *Haemophilus influenzae*. *Mol Microbiol* **36**, 784-795 (2000).
14. Hill, D.J. & Virji, M. A novel cell-binding mechanism of *Moraxella catarrhalis* ubiquitous surface protein UspA: specific targeting of the N-domain of carcinoembryonic antigen-related cell adhesion molecules by UspA1. *Mol Microbiol* **48**, 117-129 (2003).
15. Berger, C.N., Billker, O., Meyer, T.F., Servin, A.L. & 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* **52**, 963-983 (2004).
16. Barnich, N., Carvalho F. A., Glasser, A-L., Darcha, C., Jantsheff P., Allez M., Peeters H., Bommelaer, G., Desreumaux, P., Colombel, J-F., Darfeuille-Michaud, A.

- CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *J Clin Invest* **117**, 1566-74 (2007).
17. Hauck, C.R., Agerer, F., Muenzner, P. & Schmitter, T. Cellular adhesion molecules as targets for bacterial infection. *Eur J Cell Biol* **85**, 235-242 (2006).
 18. Frängsmyr, L., Baranov, V. & Hammarstrom, S. Four carcinoembryonic antigen subfamily members, CEA, NCA, BGP and CGM2, selectively expressed in the normal human colonic epithelium, are integral components of the fuzzy coat. *Tumor Biol* **20**, 277-292 (1999).
 19. Schölzel, S. et al. Carcinoembryonic Antigen Family Members CEACAM6 and CEACAM7 Are Differentially Expressed in Normal Tissues and Oppositely Deregulated in Hyperplastic Colorectal Polyps and Early Adenomas. *Am J Pathol* **156**, 595-605 (2000).
 20. Sarantis, H. & Gray-Owen, S.D. Defining the Roles of Human Carcinoembryonic Antigen-Related Cellular Adhesion Molecules during Neutrophil Responses to *Neisseria gonorrhoeae*. *Infect Immun* **80**, 345-58 (2012).
 21. Lee, H.S.W., Ostrowski, M.A. & Gray-Owen, S.D. CEACAM1 Dynamics during *Neisseria gonorrhoeae* Suppression of CD4⁺ T Lymphocyte Activation. *J Immunol* **180**, 6827-6835 (2008).
 22. Chen, Z., Chen, L., Qiao, S.-W., Nagaishi, T. & Blumberg, R.S. Carcinoembryonic Antigen-Related Cell Adhesion Molecule 1 Inhibits Proximal TCR Signaling by Targeting ZAP-70. *J Immunol* **180**, 6085-6093 (2008).

23. Hauck, C.R., Gulbins, E., Lang, F. & Meyer, T.F. Tyrosine phosphatase SHP-1 is involved in CD66-mediated phagocytosis of Opa(52)-expressing *Neisseria gonorrhoeae*. *Infect Immun* **67**, 5490-5494 (1999).
24. Sintsova, A. et al. Global Analysis of Neutrophil Responses to *Neisseria gonorrhoeae* Reveals a Self-Propagating Inflammatory Program. *Plos Pathogens* **10**, 15 (2014).
25. McCaw, S.E., Schneider, J., Liao, E.H., Zimmermann, W. & Gray-Owen, S.D. Immunoreceptor tyrosine-based activation motif phosphorylation during engulfment of *Neisseria gonorrhoeae* by the neutrophil-restricted CEACAM3 (CD66d) receptor. *Mol Microbiol* **49**, 623-637 (2003).
26. Schmitter, T., Agerer, F., Peterson, L., Münzner, P. & Hauck, C.R. Granulocyte CEACAM3 Is a Phagocytic Receptor of the Innate Immune System that Mediates Recognition and Elimination of Human-specific Pathogens. *J Exp Med* **199**, 35-46 (2004).
27. Hauck, C.R., Meyer, T.F., Lang, F. & Gulbins, E. CD66-mediated phagocytosis of Opa(52) *Neisseria gonorrhoeae* requires a Src-like tyrosine kinase- and Rac1-dependent signalling pathway. *Embo J* **17**, 443-454 (1998).
28. Johnson, M.B. et al. Opa+ *Neisseria gonorrhoeae* exhibits reduced survival in human neutrophils via Src family kinase-mediated bacterial trafficking into mature phagolysosomes. *Cell Microbiol* **17**, 648-665 (2015).
29. Fedarovich, A., J. Tomberg, R. A. Nicholas, and C. Davies. Structure of the N-terminal domain of human CEACAM1: binding target of the opacity proteins during invasion of *Neisseria meningitidis* and *N. gonorrhoeae*. *Acta Crystallogr Sec D-Biol Crystallogr* **62**, 971-979 (2006).

30. Virji, M. et al. Critical determinants of host receptor targeting by *Neisseria meningitidis* and *Neisseria gonorrhoeae* : identification of Opa adhesiotopes on the N-domain of CD66 molecules. *Mol Microbiol* **34**, 538-551 (1999).
31. Popp, A., Dehio, C., Grunert, F., Meyer, T.F. & Gray-Owen, S.D. Molecular analysis of neisserial Opa protein interactions with the CEA family of receptors: identification of determinants contributing to the differential specificities of binding. *Cell Microbiol* **1**, 169-181 (1999).
32. Kuespert, K., Weibel, S. & Hauck, C.R. Profiling of bacterial adhesin — host receptor recognition by soluble immunoglobulin superfamily domains. *Journal of Microbiological Methods* **68**, 478-485 (2007).
33. Ball, L.M. & Criss, A.K. Constitutively Opa-expressing and Opa-deficient neisseria gonorrhoeae strains differentially stimulate and survive exposure to human neutrophils. *J Bacteriol* **195**, 2982-90 (2013).
34. Fulcher, N.B. The Role of *Neisseria gonorrhoeae* Opacity Proteins in Host Cell Interactions and Pathogenesis. University of North Carolina (2004).
35. Dewald, A.H., Hodges, J. C., Columbus, L. Physical determinants of beta-barrel membrane protein folding in lipid vesicles. *Biophys J* **100**, 2131-2140 (2011).
36. Vendome, J. et al. Molecular design principles underlying β -strand swapping in the adhesive dimerization of cadherins. *Nat Struct Mol Biol* **18**, 693-700 (2011).

4. Binding affinities of a variety of Opa proteins for CEACAMs *in vitro* suggests the interaction may be non-specific

4.1 Overview

The majority of Opa proteins interact with various CEACAM receptors. While Opa – CEACAM interactions *in vivo* can result in the internalization of the Opa-expressing bacterium, the mechanisms of CEACAM mediated engulfment remain poorly defined. Gaining a deeper understanding of the Opa – CEACAM interaction may shed light on Gc and Nm engulfment mechanisms, which can potentially be exploited for therapeutic delivery. Additionally, we seek to understand the molecular determinants of Opa – CEACAM interactions. Such diverse amino acid sequences comprise the Opa HV regions that a receptor recognition motif has not been identified. Investigating the specificity and thermodynamics of Opa – CEACAM interactions is the first step to increasing our knowledge of how bacteria trigger engulfment into human host cells, and may provide insight into Opa – CEACAM molecular recognition.

Towards this end, we have expressed and folded a selection of Opa_{CEA} proteins in liposomes, as well as generated a variety of NCCM receptors to examine the Opa - CEACAM interaction using BLI. This technique determines the thermodynamic and kinetic parameters of Opa – CEACAM interactions. Additionally, the effect of varying the amount and molecular weight of PEGylated lipids in the liposomes was investigated. Different lipid:Opa protein ratios were also assessed for Opa – CEACAM interactions to determine if the amount of Opa proteins per liposome has an effect on Opa – CEACAM binding affinities. Variation of the lipid:Opa ratio will be used to determine the valency of the Opa – CEACAM interaction.

4.2 Introduction

4.2.1 *Opa_{CEA} selectivity in vivo*

Opa_{CEA} proteins differentially recognize members of the CEACAM family. To date, all experimentally investigated *Opa_{CEA}* proteins bind CEACAM1, while a subset of *Opa_{CEA}* bind various combinations of CEACAMs 3, 5, or 6.^{1,2} *Opa* proteins do not engage CEACAMs 4, 7, or 8.¹⁻³ Receptor specificity has been determined for a number of *Opa* sequences encoded for in Gc strain MS11 (specificity in Table 4.1, sequences in Table 4.2).

While a number of *Opa* variants have been classified in terms of receptor specificity, a receptor recognition motif has yet to be identified. Multiple sequence alignments of the *Opa* HV regions have not identified a motif that explains receptor selectivity. To elucidate the molecular determinants of *Opa* – CEACAM interactions, binding affinities must be determined for a variety of *Opa* proteins and CEACAM receptors. Correlation between receptor specificity and bacterial engulfment can then be investigated, which will provide insight into Neisserial pathogenicity.

Although the molecular determinants of the *Opa* – CEACAM interaction have yet to be elucidated, both HV1 and HV2 are necessary for receptor recognition.⁴ Bos *et al* investigated the binding of *Opa* proteins from Gc strain MS11, using *Opa* constructs with their extracellular loops deleted or chimeras.⁴ Loop one, which contains the SV region, is dispensable for CEACAM specificity.⁴ Deletion of either loop 2 or 3 (which contain HV1 and HV2, respectively) resulted in loss of receptor recognition.⁴

Opa mutants with loop four (the conserved loop) deleted were unable to be investigated, as this deletion resulted in a decrease in expression on the Gc surface.⁴ *Opa* chimeras, which contain sequence combinations of HV regions from a variety of *Opa_{CEA}*

variants, lost most of their receptor-binding activity (ex. An Opa chimera containing the HV1 sequence from a CEACAM1-binding Opa and the HV2 sequence from a different CEACAM1-binding Opa no longer interacts with CEACAM1).⁴ The results of this study indicate that not only are both HV1 and HV2 involved in receptor recognition, but also specific HV1 and HV2 combinations are required for CEACAM binding.⁴

Table 4.1. Receptor specificity of Opa proteins from *Neisseria gonorrhoeae* strain MS11.^{1,2}

Opa variant (rOpa)	CEACAM							HSPG
	1	3	4	5	6	7	8	
OpaA (Opa ₅₀)	-	-	-	-	-	-	-	+
OpaB (Opa ₅₇)	+	+	-	+	+	-	-	-
OpaC (Opa ₅₂)	+	+	-	+	+	-	-	-
OpaD (Opa ₅₆)	+	-	-	+	-	-	-	-
OpaE (Opa ₅₅)	+	-	-	+	-	-	-	-
OpaF (Opa ₅₄)	+	-	-	+	-	-	-	-
OpaG (Opa ₅₈)	+	+	-	+	+	-	-	-
OpaH (Opa ₅₉)	+	-	-	+	-	-	-	-
OpaI (Opa ₆₀)	+	+	-	+	+	-	-	-
OpaJ (Opa ₅₁)	+	-	-	+	-	-	-	-
OpaK (Opa ₅₃)	+	-	-	+/-	-	-	-	-

Table 4.2. HV sequences of Opa proteins from Gc strain MS11. Adapted from Bhat *et al.*⁵

rOpa	HV1 Sequence	HV2 Sequence
Opa ₅₃	S...GRDDNSTSNSS-HLNIKTQ...H.	QVR.VESE.---TTV.TH-----GAPVPQGP.PKP.Y.K.R..SSL.F.AV
Opa ₅₁--.....-----.....---KFL.SSY.-.LN....TEE...N-..HG.N.....VI
Opa ₅₀	S..NM--RVHKHNSNR.N.-----..NQI.---GLL..STP-GIM.G.YKVLRTPG-A.R..D..RRV.L.VI
Opa ₅₉	N...L.N..AN.GG.-.....R...RQQ.---IA...YPQ-NAASSVTNA.IR.LPH.E.....
Opa ₅₅	D...--ENK.Q.K-----R.L.....---.....-.....
Opa ₅₆--.....-----.....---.....-.....
Opa ₅₂	.KELERKNN.TS.GDQLNIKYQ-...H.	G.....KNTLTAYHGAGTK.TYYDDIDSGKNQKNTYR.NR.S..L.F.AM
Opa ₆₀	N.E.V--.IR.E.GI.IDR-----.....I---...VPSNAPNGAVTTYNTDPKTQNDYQ.N.....
Opa ₅₄NK.....---.....-.....
Opa ₅₈V.....-----..R.T...---.....D-.....
Opa ₅₇	SIKELLRNKGNGNR-----TDLKAENQ	SIDSTKKT---EVTTLHGPGETTPTVYPGKNTQN-AHRESDSIRRVGLGAV

. denotes identical amino acid, based off of Opa₅₇ sequences.

- denotes a gap

4.2.2 Mechanisms of CEACAM induced bacterial internalization

CEACAMs have different transmembrane domains, and cytoplasmic domains, enabling different uptake mechanisms to be triggered upon CEACAMs interacting with Opa proteins expressed on Gc. Of all the Opa-binding CEACAMs, CEACAM 1 and 3 contain a transmembrane domain, while CEACAM 5 and 6 possess a GPI anchor. Mechanisms of CEACAM mediated Gc engulfment have been described, but are not fully understood.^{2,6-10}

Engulfment via CEACAM 3 is the best understood mechanisms of Gc phagocytosis (Figure 4.1). CEACAM3 triggers Gc engulfment faster and with better efficiency than other CEACAMs.^{2,9,11,12} Upon CEACAM3 interacting with Opa, the surface of the host cell expressing CEACAM3 undergoes dramatic reorganization near the area of bacterial attachment (Figure 4.1, A, C, and E).⁷ Engulfment is initiated by F-actin filaments, which make lamellipodial protrusions up to 6 μm long that extend out and around the attached bacterium or cluster of bacteria.⁷

Billker *et al* transfected HeLa cells to express CEACAM3 to investigate internalization of Gc, and they noted that the engulfment process primarily involved the immunoreceptor tyrosine-based activation motif (ITAM) on the cytoplasmic domain of CEACAM3.⁷ Upon CEACAM3 interaction, the tyrosine residues on the ITAM are phosphorylated by the Src family kinase.⁷ Phosphorylation of the ITAM triggers the downstream activation of Rac and Cdc42 (small GTPases), which initiate the formation of the F-actin filaments, leading to the internalization of the bacterium.⁷

A more recent study found that the ITAM is not required for bacterial engulfment via CEACAM3.¹³ Sarantis *et al* used murine promyelocyte cells differentiated into neutrophils that expressed mutant CEACAM3 receptors with one or both ITAM tyrosines mutated to

phenylalanine (which is unable to be phosphorylated), or deletion mutants that lacked the ITAM or entire cytoplasmic domain.¹³ For all investigated CEACAM3 mutants, the neutrophils were able to interact with and engulf Gc that were expressing Opa proteins.¹³ However, though the mutant CEACAM3-expressing cells could engulf Gc, the neutrophils did not demonstrate the typical oxidative burst or degranulation response upon binding to Gc.¹³ These results suggest that the ITAM on the cytoplasmic domain of CEACAM3 is involved in neutrophil activation, but not Gc internalization.¹³

While CEACAM1 also contains transmembrane and cytoplasmic domains, the engulfment mechanism is quite different. Upon CEACAM1 interaction with Opa proteins on Gc, small pseudopods appear in the cell surface, which envelop the bacterium, remaining close to the surface of the cell (Figure 4.1, B, D, and F).⁷ In contrast to engulfment via CEACAM3, CEACAM1 mediated bacteria uptake does not involve massive actin rearrangement.⁷ Cholesterol rich micro-domains in the membrane play a critical role in CEACAM1 mediated engulfment, as cholesterol depletion reduces engulfment efficiency.¹⁴ Additionally, Voges *et al* found that CEACAM1 mediated bacterial uptake is dependent on PI3K, while PI3K does not affect CEACAM3 mediated engulfment.¹⁵

CEACAM1 exhibits widespread differential splicing, which leads to the creation of multiple isoforms.¹⁶ Many CEACAM1 isoforms contain an immunoreceptor tyrosine-based inhibition motif (ITIM).¹⁶ Pathogenic *Neisseria* can utilize interactions with CEACAM1 isoforms that contain ITIMs to inhibit an immune response.⁸ ITIM signaling initiated by Gc binding to CEACAM1 in CD4⁺ T cells suppressed B cell response.⁸ More recently, Slevogt *et al* found that engagement of CEACAM1 on bronchial epithelial cells by Opa expressing Nm

suppressed TLR2 signaling.¹⁷ CEACAM1 ITIM phosphorylation recruits SHP-1, which inhibits T cell signaling, and initiates negative feedback pathways by modulating PI3K.¹⁸⁻²²

The least understood bacterial engulfment mechanisms mediated by CEACAMs are via CEACAMs 5 and 6, which contain a GPI anchor instead of a transmembrane and cytoplasmic domain. GPI anchors are post-translational modifications which often direct membrane proteins to membrane micro-domains enhanced with cholesterol and sphingolipids.²³ Upon CEACAM activation, engulfment appears to occur via a “zipper-like” mechanism, where neighboring CEACAM receptors are progressively recruited because of a large amount of Opa protein on a bacterium.^{9,10} Eventually, the bacterium will be fully surrounded by the host cell membrane, which “zippered” together around it. As GPI anchored CEACAMs lack the cytoplasmic domain, they are unable to be phosphorylated like CEACAMs 1 and 3.¹⁴ However, the GPI anchor is important to trigger bacterial engulfment, as cleavage of the anchor results in a significant decrease of bacteria internalization.⁹

GPI anchored proteins have been shown to be involved in cell signaling as well. Typically, GPI anchored proteins are able to mediate intracellular signaling events after antibody-induced ligation.²⁴ Hiscox *et al* determined that GPI anchored CD59 receptors were able to contribute to Ca^{2+} signaling.²⁴ However, the GPI anchor was not directly involved in signaling; rather it is clustering of the GPI anchored proteins that is necessary for intracellular signaling.²⁴ In fact, not only do GPI anchored proteins play a role in $[\text{Ca}^{2+}]$ modulation, but have been shown to be involved in tyrosine phosphorylation of intracellular proteins, inositol phosphate turnover, and cytokine secretion as well.²⁵⁻³⁰ Therefore, we can speculate that bacterial engulfment via CEACAM5 or 6 may be dependent on clustering of the receptors,

which could be initiated by binding to Opa proteins. Once the CEACAM receptors have clustered, intracellular signaling pathways are triggered that induces bacterial uptake.

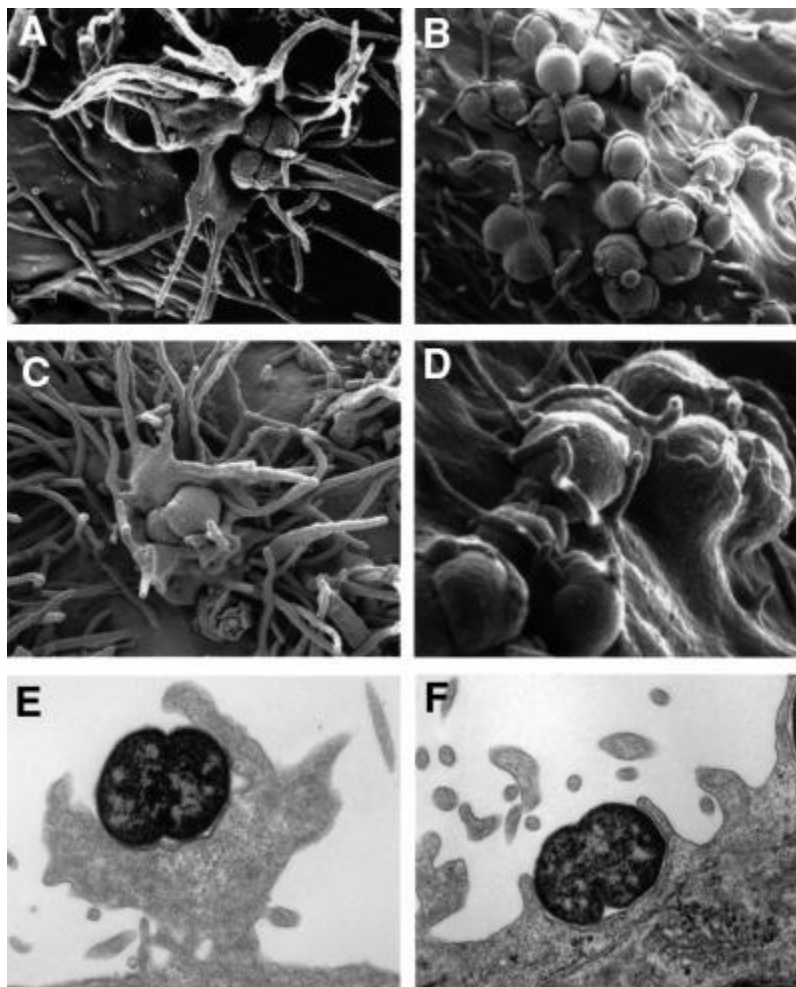


Figure 4.1. *Neisseria gonorrhoeae* N313 (Opa₅₇) triggers different structural modifications on HeLa cell surfaces, depending on the expression of CEACAM1 or CEACAM3. *Reprinted with permission.*⁷ Scanning electron micrographs of HeLa cells infected with Gc show large protrusions when CEACAM3 is expressed (A and C), or pseudopods remaining close to cell surface when CEACAM1 is expressed (B and D). Tunneling electron micrographs of HeLa cells engulfing Gc via CEACAM3 (E) or CEACAM1 (F).

4.3 Results and Discussion

4.3.1 *The Opa – CEACAM interaction is of high affinity for all combinations investigated*

The results of the FP experiments described in chapter 3 indicated that interactions between Opa₆₀ or OpaD and NCCM1 or NCCM3 are of high affinity (nM). Additional binding experiments using biolayer interferometry (BLI) were conducted to verify these tight affinities, as well as to investigate additional Opa protein variants with different CEACAM receptors (representative data shown in Figure 4.2). The theory behind BLI is detailed in chapter 2.

A variety of Opa_{CEA} proteins (Opa₆₀ from Gc strain MS11, OpaD and OpaI from Gc strain FA1090) and non-Opa_{CEA} proteins (Opa₅₀ from Gc strain MS11, OpaA from Gc strain FA1090) were assessed for binding to different CEACAMs (NCCM1, 3, 4, 5, and 8). CEACAMs 1, 3, and 5 bind differentially to Opa_{CEA} proteins, while CEACAMs 4 and 8 do not interact with any Opa proteins investigated to date.⁶ All Opa proteins were folded and reconstituted into the final lipid mixture described in chapter 2. GST-NCCM constructs were obtained or created for all CEACAM N-domains mentioned above. The GST-NCCM was immobilized onto a biosensor coated with anti-GST antibodies, washed, and incubated with the Opa proteoliposomes.

The affinities of the Opa_{CEA} proteins for the CEACAM receptors will be presented in the following order: NCCM1, NCCM3, NCCM4, NCCM5, and NCCM8 (Table 4.3). Affinities of Opa₆₀ are calculated to be 20 ± 4 nM, 28 ± 12 nM, 16 ± 5 nM, 29 ± 12 nM, and 24 ± 6 nM. OpaD affinities are 30 ± 5 nM, 35 ± 19 nM, 19 ± 6 nM, 31 ± 7 nM, and 30 ± 5 nM. Finally, affinities of OpaI are calculated to be 22 ± 5 nM, 31 ± 22 nM, 16 ± 6 nM, 23 ± 3 nM, 22 ± 2 nM.

For the non-Op_{ACEA} proteins, affinities of Opa₅₀ for NCCM1, 3, 4, 5, and 8 were calculated as 30 ± 2 nM, 29 ± 6 nM, 18 ± 5 nM, 28 ± 8 nM, and 23 ± 7 nM, respectively (Table 4.4). OpaA demonstrated affinities of 19 ± 1 nM for NCCM1, 39 ± 37 nM for NCCM3, 17 ± 8 for NCCM4, 29 ± 5 nM for NCCM5, and 25 ± 4 nM for NCCM8 (Table 4.4). Empty liposomes were also investigated for interactions with all NCCMs, but an association curve could not be generated, as binding levels were too low to detect (Figure 4.2, blue). Therefore, all Opa binding curves are background subtracted using the liposomes as a reference. OpaA proteins reconstituted into DMPC liposomes were used as a negative control, for which no association was detected for all NCCMs (data not shown).

Op_{ACEA} proteins reconstituted *in vitro* do not show CEACAM receptor binding specificity. Opa₆₀ binds to NCCMs 1, 3, 4, 5, and 8 with similar affinities (Table 4.3). The same can be said for OpaD, which has similar K_D values for interactions with all NCCMs (Table 4.3). OpaI also demonstrates equal binding affinity towards all NCCM receptors (Table 4.3). Non-Op_{ACEA} proteins also demonstrate this lack of CEACAM receptor specificity. Opa₅₀ interactions with all NCCMs are comparable to one another (Table 4.4). The affinities of OpaA for the NCCM receptors are all similar as well (Table 4.4).

While CEACAMs 4 and 8 have previously been shown to not interact with Opa proteins *in vivo*,⁶ they tightly bind Opa proteoliposomes *in vitro* (Table 4.3, Table 4.4). NCCMs 1, 3, 4, 5, and 8 all display approximately equal nanomolar affinity towards all Op_{ACEA} proteins (Opa₆₀, OpaD, and OpaI, Table 4.3), as well as towards the non-Op_{ACEA} proteins (Opa₅₀ and OpaA, Table 4.4). Most of the NCCM interactions with Opa proteins demonstrated low error across triplicates; however, NCCM3 interactions were quite variable for all Opa proteins.

Affinity constants for Opa₆₀ and OpaD with NCCM1 and NCCM3 were determined using fluorescence polarization (see Chapter 3). K_D values determined via FP for both Opa₆₀ and OpaD are approximately 2 nM for NCCM1, and 5 nM for NCCM3. Comparing these K_D values obtained from FP to those calculated using BLI (~20 nM for both Opa₆₀ and OpaD with NCCM1 and NCCM3), the affinities are quite similar. Both techniques yielded data which supports the claim that Opa - CEACAM interactions are high affinity.

An additional benefit to using BLI over FP is that BLI provides kinetic information about interactions in addition to binding affinities. Association constants (k_{on}) and dissociation constants (k_{off}) for all Opa proteins binding to all NCCMs are reported in tables 4.3 and 4.4. K_D is the ratio of k_{off}/k_{on} , so high affinity interactions, such as those calculated for Opa – NCCM binding, typically have fast on rates and slow off rates. This is the case for all Opa – NCCM interactions investigated. Association constants range from $5.2 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for the OpaD – NCCM5 interaction, to $1.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ for the OpaI – NCCM4 interaction (Table 4.3). This corresponds to Opa – NCCM association rates between 0.83 μs – 1.9 μs .

Dissociation rates for all Opa – NCCM interactions are similar to one another, and on the order of minutes. The fastest dissociation rate observed was for the interaction of Opa₅₀ and NCCM1, with a k_{off} value of $2 \times 10^{-2} \text{ s}^{-1}$ (Table 4.4). The slowest dissociation rate was observed between Opa₆₀ and NCCM8 as well as between Opa₅₀ and NCCM8, with a k_{off} equal to $1.5 \times 10^{-2} \text{ s}^{-1}$ (Table 4.3). These k_{off} values correspond to dissociation rates ranging between 50 s – 71 s.

While little is known about dissociation rates for membrane protein interactions, likely due to the difficulty in developing a functional *in vitro* system, many other protein – protein dissociation rates have been investigated. Direct interaction of recombinant G $_{\beta 1\gamma 2}$ with the C-

terminal domain of Kir3.4 (a G protein-gated inward rectifier K⁺ channel subunit) was found to have a dissociation rate of $\sim 0.003 \text{ s}^{-1}$ (333 s).³¹ Svensson *et al* investigated the binding of PLW (Y47W mutant of the B1 domain of Protein L) to the Igk light chain, and found a fast dissociation phase and a slow dissociation phase.¹⁰ The fast phase dissociation rate was $8.7 \times 10^{-2} \text{ s}^{-1}$ ($\sim 11.5 \text{ s}$), which is approximately 5 times faster than the Opa – NCCM dissociation rate. The slow phase dissociation rate was $6.3 \times 10^{-3} \text{ s}^{-1}$ ($\sim 159 \text{ s}$), which is about twice as slow as the Opa – NCCM dissociation rate.³²

Zhang *et al* determined the dissociation rates of the dimerization β LG-A to be 0.52 s^{-1} and 0.37 s^{-1} (2-3 s) for the N- and C-terminal interfaces respectively, which is much faster than what we observe for Opa – NCCM interactions.³³ The dissociation rate for a heterotypic interaction between the spore coat proteins CotY and CotZ from *Bacillus subtilis* was calculated to be 1.3 s^{-1} ($\sim 0.8 \text{ s}$).³⁴ Examination of a cytokine receptor complex involving the extracellular domains of type 1 interferon receptors with the IFN α 2 ligand yielded a dissociation rate of 0.02 s^{-1} (50 s).³⁵

CD2, a Ig cell-adhesion molecule, was found to have a very low affinity (60-90 μM) for CD48, with a dissociation rate of 6 s^{-1} ($\sim 0.17 \text{ s}$).³⁶ This study utilized receptor proteins from rats, while investigations with recombinant human CD2 and CD58 receptors reported a higher affinity (0.4 μM).³⁷ A separate investigation using recombinant human CD2 and human CD58 found the interaction to have an affinity $\sim 9\text{-}22 \text{ }\mu\text{M}$ and a $k_{\text{off}} \geq 4 \text{ s}^{-1}$ (0.25 s).³⁸ These affinities for CD2 are much lower than those calculated for Opa – NCCM interactions, and the differences can be explained by the fast k_{off} rates for CD2 compared to the much slower dissociation of Opa – NCCM.

Comparing the calculated rates for the Opa – NCCM dissociation to these previously published rates, the Opa – NCCM dissociation rate is much slower than some rates but approximately equal to or faster than other dissociation rates. This suggests that the dissociation values obtained from BLI for Opa – NCCM interactions are reasonable, as they are comparable to what is known about other high affinity interactions.

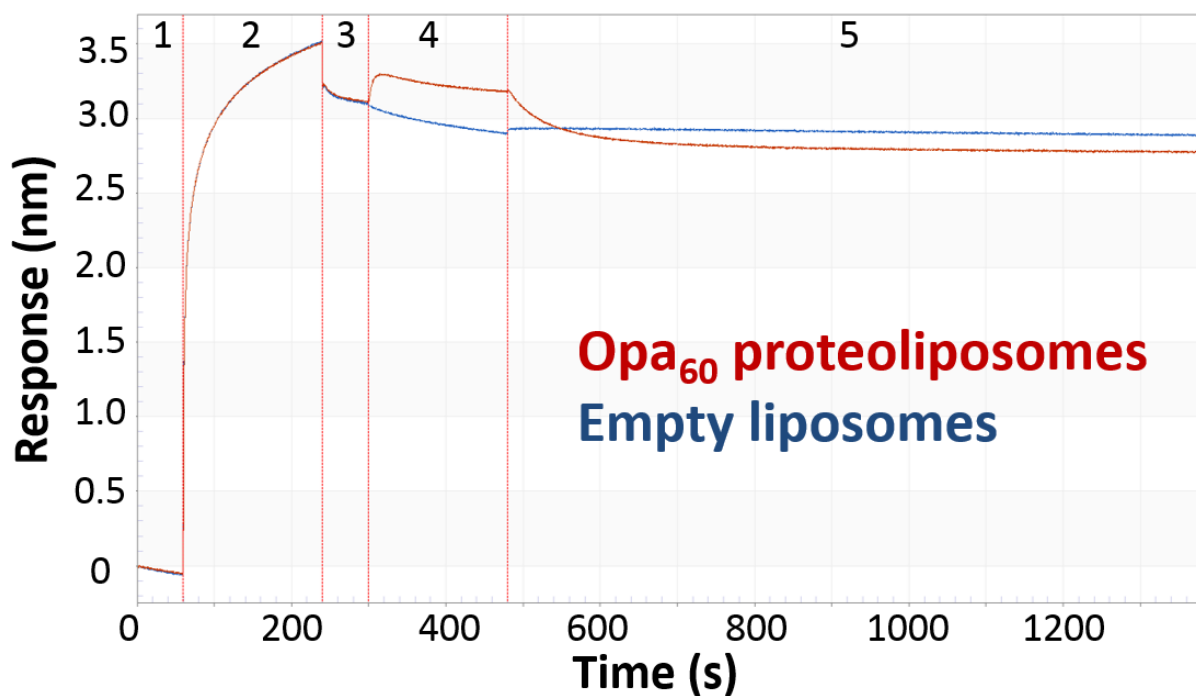


Figure 4.2. Representative raw BLI binding data. GST biosensors were dipped into buffer (panel 1), and then incubated with NCCM1 (panel 2). The biosensors were dipped into buffer again to remove weakly associated NCCM1 (panel 3). Panel 4 shows the association of Opa₆₀ proteoliposomes (red) or empty liposomes (blue) with the immobilized GST-NCCM1 on the surface of the biosensor. Dissociation is recorded in panel 5.

Table 4.3. Kinetic and thermodynamic parameters for Opa_{CEA} interactions with a variety of NCCM receptors. Error is the standard deviation of samples in triplicate.

Opa _{CEA}	Opa ₆₀			Opa _D			Opa _I		
	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)
NCCM1	9.4 ± 3.2	1.7 ± 0.2	20 ± 4	6.1 ± 0.3	1.8 ± 0.2	30 ± 5	8.9 ± 1.2	1.9 ± 0.1	22 ± 5
NCCM3	7.0 ± 3.2	1.6 ± 0.4	28 ± 12	6.4 ± 3.1	1.6 ± 0.5	35 ± 19	9.4 ± 5.1	1.7 ± 0.3	31 ± 22
NCCM4	10 ± 1.9	1.6 ± 0.3	16 ± 5	9.3 ± 3	1.6 ± 0.2	19 ± 6	12 ± 5	1.6 ± 0.3	16 ± 6
NCCM5	6.8 ± 3.0	1.6 ± 0.2	29 ± 12	5.2 ± 1.0	1.6 ± 0.1	31 ± 7	7.0 ± 0.3	1.6 ± 0.2	23 ± 3
NCCM8	7.0 ± 2.0	1.5 ± 0.1	24 ± 6	5.3 ± 0.7	1.6 ± 0.1	30 ± 5	7.5 ± 0.8	1.6 ± 0.1	22 ± 2

Table 4.4 Kinetic and thermodynamic parameters for non-Opa_{CEA} proteins interacting with a variety of NCCM receptors. Error is the standard deviation of samples in triplicate.

Non-Opa _{CEA}	Opa ₅₀			OpaA		
	$k_{on} \times 10^5$ (M ⁻¹ s ⁻¹)	$k_{off} \times 10^{-2}$ (s ⁻¹)	K _D (nM)	$k_{on} \times 10^5$ (M ⁻¹ s ⁻¹)	$k_{off} \times 10^{-2}$ (s ⁻¹)	K _D (nM)
NCCM1	6.8 ± 1.1	2.0 ± 0.2	30 ± 2	8.9 ± 0.5	1.7 ± 0.1	19 ± 1
NCCM3	6.6 ± 2.3	1.8 ± 0.3	29 ± 6	9.1 ± 5.4	1.6 ± 0.2	39 ± 37
NCCM4	9.4 ± 1.5	1.6 ± 0.1	18 ± 5	11 ± 5	1.4 ± 0.2	17 ± 8
NCCM5	6.4 ± 1.9	1.6 ± 0.1	28 ± 8	6.0 ± 0.9	1.7 ± 0.3	29 ± 5
NCCM8	7.4 ± 2.6	1.5 ± 0.1	23 ± 7	6.7 ± 1.0	1.6 ± 0.1	25 ± 4

4.3.2 Varying the amount or molecular weight of PEGylated lipids has little to no effect on Opa – CEACAM interactions

As discussed in chapter 2, the addition of PEGylated lipid to liposomes can change the behavior of the liposome. How much PEGylated lipid is incorporated as well as the molecular weight of the PEG polymer can play a role in modulating the function of the liposomes.³⁹ Du *et al* and Price *et al* both noted that the amount and molecular weight of the PEGylated lipids had an effect on protein adsorption, as well as cell adhesion.^{40,41}

To determine if the amount of PEGylated lipid or molecular weight of the PEG polymer affects the Opa – CEACAM interaction, BLI experiments were conducted using GST-NCCM1 with Anti-GST biosensors to interact with Opa₆₀ proteins folded in different lipid environments. Varying amounts of PEG1000 (5 mol%, 10 mol%, or 15 mol%) or PEG2000 (5 mol%, 10 mol%, or 15 mol%) were incorporated into liposomes composed of DMPC, DMPG, and cholesterol (described in chapter 2) to determine how they affected the Opa₆₀ – NCCM1 interaction.

Binding affinities of Opa₆₀ in liposomes with PEG1000 incorporated were similar regardless of how much PEG1000 was present. 5 mol% yielded a K_D of 19.7 nM, compared to 37.5 nM for 10 mol% and 32 nM for 15 mol% (Table 4.5). Incorporation of PEG2000 yielded the same trend, with dissociation constants of 8.2 nM for 5 mol%, 6.7 nM for 10 mol%, and 6.8 nM for 15 mol%, which are all similar (Table 4.5). Opa₆₀ – NCCM1 affinities with PEG2000 are approximately one half to one fourth of those with PEG1000, but with such tight affinities these differences may not be important.

Association rates of Opa₆₀ and NCCM1 were similar as well, regardless of the molecular weight or amount of PEG present on the liposomes. PEG1000 liposomes

demonstrated k_{on} values of $9.4 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for 5 mol%, $5.6 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for 10 mol%, and $6.5 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ for 15 mol% (Table 4.5, top). These k_{on} values represent association rates ranging from $1.1 \mu\text{s} - 1.8 \mu\text{s}$. For PEG2000, k_{on} values of $1.4 \times 10^6 \text{ s}^{-1}$, $1.8 \times 10^6 \text{ s}^{-1}$, and $2.0 \times 10^6 \text{ s}^{-1}$ were determined for 5 mol%, 10 mol%, and 15 mol%, respectively, which yields association rates between $0.5 \mu\text{s} - 0.7 \mu\text{s}$ (Table 4.5, bottom).

Dissociation of Opa₆₀ and NCCM1 also occurred at the same rate, regardless of the amount or molecular weight of the PEG in the liposomes. PEG1000 dissociation constants are $2.28 \times 10^{-2} \text{ s}^{-1}$, $2.11 \times 10^{-2} \text{ s}^{-1}$, and $2.07 \times 10^{-2} \text{ s}^{-1}$ for 5 mol%, 10 mol%, and 15 mol% respectively (Table 4.5, top). These k_{off} values correspond to 44 – 48 s for dissociation. The dissociation rate values for PEG2000 are $1.1 \times 10^{-2} \text{ s}^{-1}$ for 5 mol%, $1.2 \times 10^{-2} \text{ s}^{-1}$ for 10 mol%, and $1.4 \times 10^{-2} \text{ s}^{-1}$ for 15 mol%, yielding a range of 71 – 91 s for the dissociation rates. Again, while there are slight differences between the k_{on} and k_{off} values between PEG1000 and PEG2000, as well as between the different mol% incorporated into the liposomes, the differences are not substantial.

Table 4.5. Kinetic and thermodynamic parameters for the Opa60 – NCCM1 interaction with varied amounts of PEG1000 (top) or PEG2000 (bottom) in the liposomes. Error is the standard deviation of samples in triplicate.

PEG1000	5 Mol%			10 Mol%			15 Mol%		
	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)
NCCM1	9.4	1.7	19.7	5.64	2.11	37.5	6.46	2.07	32.0
PEG2000	5 Mol%			10 Mol%			15 Mol%		
	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)
NCCM1	14 ± 0.7	1.1 ± 0.004	8.2 ± 0.4	18 ± 1.6	1.2 ± 0.04	6.8 ± 0.6	20 ± 3	1.4 ± 0.16	6.8 ± 1.2

4.3.3 The lipid:Opa ratio has little effect on Opa association with NCCM1.

Opa₆₀ was reconstituted into varying concentrations of lipid (liposome composition is the same as described in chapter 2) to determine if the lipid:Opa protein ratio affects the Opa – CEACAM interaction. BLI experiments were carried out using Anti-GST biosensors with GST-NCCM1 immobilized for interaction with the Opa proteoliposomes. Previous binding experiments described above were conducted with a lipid:Opa ratio of 234:1, so several ratios above (350:1 467:1) and below (111:1 and 70:1) were chosen for investigations.

All lipid:Opa ratios demonstrated similar Opa – NCCM1 affinities (Table 4.6, data for 111:1 not shown). The lowest lipid:Opa ratio of 70:1 demonstrated the weakest affinity between Opa₆₀ and NCCM1, with a K_D of 44 nM (Table 4.6). Ratios of 111:1, 350:1, and 467:1 demonstrated K_D values of 19 nM, 22 nM, and 27 nM, respectively, all of which are approximately twice as tight as at the Opa₆₀ – NCCM1 interaction with a lipid:Opa ratio of 70:1 (Table 4.6).

Association rates of Opa₆₀ and NCCM1 with lipid:Opa ratios of 70:1, 111:1, 350:1, and 467:1 are $2.8 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$, $8.1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$, $5.1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$, and $4.1 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$, respectively. Association with a lipid:Opa ratio of 70:1 was the slowest, with a rate of 3.6 μs . All other association rates ranges between 1.3 – 2.4 μs . While the k_{on} values for association of Opa₆₀ with NCCM1 were variable with different lipid:Opa ratios, the dissociation rates were very similar. A ratio of 70:1 yielded a dissociation rate of $1.2 \times 10^{-2} \text{ s}^{-1}$. Changing the ratio to 111:1 resulted in a k_{off} value of $1.4 \times 10^{-2} \text{ s}^{-1}$. Ratios of 351:1 and 467:1 yielded the same k_{off} value, $1.1 \times 10^{-2} \text{ s}^{-1}$. These k_{off} values equal dissociation rates between 71 – 91 s.

Changing the lipid:Opa ratio modulates the size of the liposome, which must also be taken into account. Increasing the lipid:Opa ratio increases the size of the liposomes (Table

4.7). Opa proteoliposomes with a lipid:Opa ratio of 70:1 exhibited a radius of 44.5 nm, with 30.1% polydispersity (Table 4.7). A lipid:Opa ratio of 350:1 yielded liposomes with an average radius of 66.4 nm, and a polydispersity of 29.6% (Table 4.7). The highest lipid:Opa ratio investigated, 467:1, produced liposomes with an average radius of 85.7 nm with 28.9% polydispersity (Table 4.7). While these sizes are variable, they all still are representative of SUVs (which can be up to 100 nm in radius).⁴²

Table 4.6. Kinetic and thermodynamic parameters for the Opa⁶⁰ – NCCM1 interaction with varied lipid:Opa. Error is the standard deviation of samples in triplicate.

Lipid:Opa	70:1			350:1			467:1		
	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)	$k_{on} \times 10^5$ ($M^{-1}s^{-1}$)	$k_{off} \times 10^{-2}$ (s^{-1})	K_D (nM)
NCCM1	2.8 ± 0.1	1.2 ± 0.03	44 ± 2	5.1 ± 0.1	1.1 ± 0.01	22 ± 0.5	4.13 ± 0.01	1.1 ± 0.01	27 ± 0.2

Table 4.7 Liposome size and polydispersity with various lipid:Opa ratios.

Lipid:Opa	70:1	350:1	467:1
Radius (nm)	44.5	66.4	85.7
Polydispersity	30.1 %	29.6 %	28.9 %

4.3.4 Understanding Opa – CEACAM selectivity

Affinities calculated for the Opa₆₀ and OpaD proteoliposome interactions with NCCM1 or NCCM3 using BLI were very similar to those observed in FP experiments (presented in chapter 3). This further supports that Opa – CEACAM interactions are of extremely tight affinities. However, affinities of the non-Opa_{CEA} proteins for all NCCMs are comparable. Additionally, affinities of all Opa_{CEA} proteins for all NCCMs are similar, which suggests that for our *in vitro* system, the Opa – CEACAM interaction is non-specific.

Several possible scenarios may be able to explain this lack of specificity in our *in vitro* Opa proteoliposome system. Interactions between CEACAM and pathogenic *Neisseria* allow the bacteria to adhere to and colonize human cells, but can also trigger engulfment of the bacteria.^{6,43} CEACAMs 1, 3, 5, and 6 have been previously shown to trigger engulfment of bacterial pathogens, while CEACAMs 4 and 8 do not.^{6,9,44-48} Conversely, our calculated affinities suggest that Opa proteins bind to CEACAMs non-preferentially. However, our *in vitro* system contains only purified recombinant Opa proteins, without any other surface structures present on Gc *in vivo*. While this is necessary to study the effects due to Opa proteins alone, Opa – receptor selectivity may depend on other cellular factors not present in our proteoliposome system.

Neisseria possess LOS on their outer membrane (described in chapter 2). LOS extends above the cell membrane, which may spatially restrict the motion of the Opa extracellular loops. Interactions between Opa proteins and LOS have been demonstrated, and are purported to be electrostatic in nature and involve the basic amino acid residues on the Opa extracellular loops.⁴⁹⁻⁵¹ PEGylated lipids were incorporated into the Opa proteoliposome to mimic the spatial restrictions due to LOS, but the PEG polymer will not directly interact with the Opa

extracellular loops. Additionally, the amount of PEGylated lipid in the Opa proteoliposomes and the size of the PEG polymers did not have an effect on Opa₆₀ – NCCM1 interactions.

We can theorize that these electrostatic interactions between LOS and Opa loops may play a large role in conformation sampling of the Opa loops, and loss of these electrostatic interactions abolishes receptor specificity. Pull down experiments (see Chapter 3) demonstrated that when expressed in Gc, Opa₅₀ was minimally interacting with recombinant NCCM1 and NCCM3, while recombinant Opa₅₀ proteoliposomes interacted with both NCCMs to a higher extent.

An alternative explanation for this observed lack of specificity of Opa – CEACAM interactions *in vitro* may be that Opa proteins are more promiscuous in interacting with receptors. Perhaps Opa proteins are able to interact with many CEACAM receptors, but these interactions are unable to trigger engulfment of the bacteria. This scenario would also provide a better understanding of how Opa sequences can be so diverse and yet still engage such a small subset of receptors. Invasion assays with Opa proteoliposomes are currently being investigated by Jason Kuhn of the Columbus laboratory, in collaboration with the Criss laboratory, to determine if our calculated *in vitro* affinities of the Opa – CEACAM interaction correlate to invasion efficacy. In these experiments, CEACAM1 receptor mediated engulfment of Opa₆₀ proteoliposomes is observed (Figure 4.3).

As CEACAM1 mediated engulfment is seen by Jason Kuhn using the Opa₆₀ proteoliposomes, this may suggest that the isolated NCCM domain is not sufficient for bacterial internalization. In fact, Voges *et al* observed a dependence of Gc engulfment on the presence of the extracellular IgC2 domain of CEACAM1.¹⁵ Binding specificities of Opa proteins from Gc strain MS11, presented in Table 4.2, were experimentally determined using

invasion assays with Opa expressing Gc. Adhesion experiments should be utilized along with invasion assays to determine if these Opa proteins are able to bind to and adhere to cells without being engulfed. Taken together, we can hypothesize that Opa proteins are able to interact with the N-terminal domain of CEACAMs non-specifically, but engulfment of the bacteria is more selective, and requires host factors beyond the NCCM domain alone. This may explain why such variable Opa protein HV1 and HV2 sequences are able to engage such a small selection of receptors.

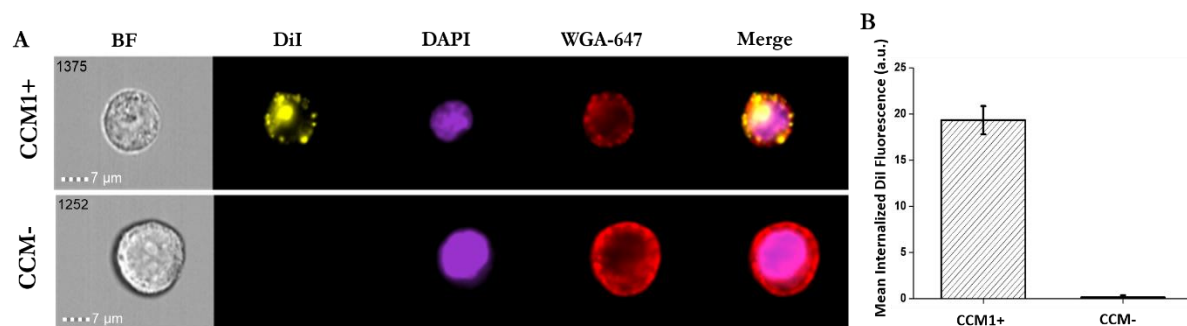


Figure 4.3. CEACAM1+ HeLa cells show higher internalization of Opa₆₀ proteoliposomes than CEACAM- cells. Image courtesy of Jason Kuhn. Opa₆₀ proteoliposomes were incubated with CEACAM1 expressing (CEACAM+) HeLa cells and HeLa cells not transfected to express CEACAM1 (CEACAM-). Cells were imaged using fluorescence microscopy (A), and the mean internalized fluorescence was calculated (B). Opa₆₀ proteoliposomes are labeled with DiI, DAPI stains the nucleus of the cell, and WGA-647 stains the cell membrane (A). Error bars in (B) represent a 99% confidence interval.

4.3.5 *Opa – CEACAM interactions appear to be monovalent*

Affinities of Opa₆₀ – NCCM1 interactions with Opa₆₀ folded in various lipid:Opa ratios can be used as a reporter to investigate whether the Opa – CEACAM interaction is multivalent. Multivalent interactions occur when more than one molecular recognition event occurs simultaneously (ex. two ligands bind a protein).⁵² If the Opa – CEACAM interaction is multivalent, binding affinities would be different for various lipid:Opa ratios, while if the Opa – CEACAM interaction is monovalent, the lipid:Opa ratio would have no effect on K_D .

With multivalent interactions, each binding event increases the likelihood that other binding events will occur.⁵³ Multivalent interactions are beneficial in drug design, as they are able to convert inhibitors with low affinity (on the mM or μ M order) to high avidity (accumulated strength of multiple individual affinities), which can improve the potency of a drug.⁵⁴ Antibodies are prime examples of proteins with multivalent interactions, as antibodies have multiple equivalent binding sites.⁵⁵ On cell surfaces, monovalent receptors and ligands can sometimes function as a multivalent system.⁵⁶

Recall from chapter 2 that K_D is equal to the concentration of reactants over the concentration of products (Eq. 2.2). This equation assumes that one receptor is binding to one ligand. When you have more than one ligand binding to a receptor, the multivalent affinity constant (K_N) is calculated using the following equation:

$$K_N = \frac{[RL_N]}{[R][L]^N} \quad \text{Eq. 4.1}$$

Where R is receptor, L is ligand, and N is equal to the number of ligand binding sites on the receptor.⁵⁷

The K_D values calculated for the Opa₆₀ – NCCM1 interaction with all of the various lipid:Opa protein ratios are similar. This suggests that the Opa – CEACAM interaction is not

multivalent. Interestingly, the lipid:Opa ratio yielded a K_D value approximately double the values for the rest of the ratios investigated. However, the highest polydispersity was seen in the proteoliposomes with a ratio of 70:1, and could be indicative that this ratio is too low to form stable SUVs.

4.4 Concluding remarks

Binding affinities for a selection of Opa proteins (Opa₆₀, OpaD, OpaI, Opa₅₀, and Opa A) with various NCCM receptors (NCCM1, NCCM3, NCCM4, NCCM5, and NCCM8) were assessed using BLI. All investigated interactions were of a similarly tight affinity, suggesting the Opa – NCCM interaction *in vitro* is non-specific. Various amounts and molecular weights of PEGylated lipids were investigated for their effect on the Opa₆₀ – NCCM1 interaction. Neither the molecular weight nor the amount of PEGylated lipid incorporated into the liposome had an effect on the Opa₆₀ – NCCM1 interaction. Additionally, the lipid:Opa ratio was varied to determine the impact on the Opa₆₀ – NCCM1 interaction. Binding affinities for all investigated lipid:Opa ratios were similar, which suggests that the Opa – CEACAM interaction is monovalent.

4.5 References

1. Bos, M.P., Hogan, D. & Belland, R.J. Selection of Opa+ *Neisseria gonorrhoeae* by limited availability of normal human serum. *Infect Immun* **65**, 645-50 (1997).
2. Gray-Owen, S.D., Lorenzen, D.R., Haude, A., Meyer, T.F. & Dehio, C. Differential Opa specificities for CD66 receptors influence tissue interactions and cellular response to *Neisseria gonorrhoeae*. *Mol Microbiol* **26**, 971-80 (1997).
3. Fulcher, N.B. University of North Carolina (2004).
4. Bos, M.P., Kao, D., Hogan, D.M., Grant, C.C. & Belland, R.J. Carcinoembryonic antigen family receptor recognition by gonococcal Opa proteins requires distinct combinations of hypervariable Opa protein domains. *Infect Immun* **70**, 1715-23 (2002).
5. Bhat, K.S. et al. The opacity proteins of *Neisseria gonorrhoeae* strain MS11 are encoded by a family of 11 complete genes. *Mol Microbiol* **5**, 1889-1901 (1991).
6. Kuespert, K., Pils, S. & Hauck, C.R. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol* **18**, 565-571 (2006).
7. Billker, O. et al. Distinct mechanisms of internalization of *Neisseria gonorrhoeae* by members of the CEACAM receptor family involving Rac1- and Cdc42-dependent and -independent pathways. *Embo J* **21**, 560-71 (2002).
8. Boulton, I.C. & Gray-Owen, S.D. Neisserial binding to CEACAM1 arrests the activation and proliferation of CD4+ T lymphocytes. *Nat Immunol* **3**, 229-36 (2002).
9. McCaw, S.E., Liao, E.H. & Gray-Owen, S.D. Engulfment of *Neisseria gonorrhoeae*: Revealing Distinct Processes of Bacterial Entry by Individual Carcinoembryonic

- Antigen-Related Cellular Adhesion Molecule Family Receptors. *Infect Immun* **72**, 2742-2752 (2004).
10. Swanson, J.A. & Baer, S.C. Phagocytosis by zippers and triggers. *Trends Cell Biol* **5**, 89-93 (1995).
 11. Chen, T., Grunert, F., Medina-Marino, A. & Gotschlich, E.C. Several Carcinoembryonic Antigens (CD66) Serve as Receptors for Gonococcal Opacity Proteins. *J Exp Med* **185**, 1557-1564 (1997).
 12. Bos, M.P., Grunert, F. & Belland, R.J. Differential recognition of members of the carcinoembryonic antigen family by Opa variants of *Neisseria gonorrhoeae*. *Infect Immun* **65**, 2353-61 (1997).
 13. Sarantis, H. & Gray-Owen, S.D. Defining the Roles of Human Carcinoembryonic Antigen-Related Cellular Adhesion Molecules during Neutrophil Responses to *Neisseria gonorrhoeae*. *Infect Immun* **80**, 345-58 (2012).
 14. Schmitter, T. et al. Opa Proteins of Pathogenic *Neisseriae* Initiate Src Kinase-Dependent or Lipid Raft-Mediated Uptake via Distinct Human Carcinoembryonic Antigen-Related Cell Adhesion Molecule Isoforms. *Infect Immun* **75**, 4116-4126 (2007).
 15. Voges, M., Bachmann, V., Naujoks, J., Kopp, K. & Hauck, C.R. Extracellular IgC2 constant domains of CEACAMs mediate PI3K sensitivity during uptake of pathogens. *PLoS One* **7**, e39908 (2012).
 16. Kammerer, R. & Zimmermann, W. Coevolution of activating and inhibitory receptors within mammalian carcinoembryonic antigen families. *BMC Biology* **8**, 12-12 (2010).

17. Slevogt, H. et al. CEACAM1 inhibits Toll-like receptor 2-triggered antibacterial responses of human pulmonary epithelial cells. *Nat Immunol* **9**, 1270-1278 (2008).
18. Barrow, A.D. & Trowsdale, J. You say ITAM and I say ITIM, let's call the whole thing off: the ambiguity of immunoreceptor signalling. *Eur J Immunol* **36**, 1646-1653 (2006).
19. Nagaishi, T. et al. SHP1 Phosphatase-Dependent T Cell Inhibition by CEACAM1 Adhesion Molecule Isoforms. *Immunity* **25**, 769-781 (2006).
20. Ravetch, J.V. & Lanier, L.L. Immune inhibitory receptors. *Science* **290**, 84-9 (2000).
21. Bousquet, C. et al. Direct binding of p85 to sst2 somatostatin receptor reveals a novel mechanism for inhibiting PI3K pathway. *Embo J* **25**, 3943-54 (2006).
22. Dubois, M.J. et al. The SHP-1 protein tyrosine phosphatase negatively modulates glucose homeostasis. *Nat Med* **12**, 549-56 (2006).
23. Simons, K. & Toomre, D. Lipid rafts and signal transduction. *Nat Rev Mol Cell Biol* **1**, 31-9 (2000).
24. Hiscox, S., Hallett, M.B., Paul Morgan, B. & van den Berg, C.W. GPI-anchored GFP signals Ca^{2+} but is homogeneously distributed on the cell surface. *Biochem Biophys Res Commun* **293**, 714-721 (2002).
25. Robinson, P.J. Glycosyl-Phosphatidyl-Inositol Membrane Anchors Signal transduction by GPI-anchored membrane proteins. *Cell Biol Int Rep* **15**, 761-767 (1991).
26. Hořejší, V. et al. Signal transduction in leucocytes via GPI-anchored proteins: an experimental artefact or an aspect of immunoreceptor function? *Immunol Lett* **63**, 63-73 (1998).

27. Deckert, M., Ticchioni, M., Mari, B., Mary, D. & Bernard, A. The glycosylphosphatidylinositol-anchored CD59 protein stimulates both T cell receptor ζ /ZAP-70-dependent and -independent signaling pathways in T cells. *Eur J Immunology* **25**, 1815-1822 (1995).
28. Cebecauer, M., Černý, J. & Hořejší, V. Incorporation of Leucocyte GPI-Anchored Proteins and Protein Tyrosine Kinases into Lipid-Rich Membrane Domains of COS-7 Cells. *Biochem Biophys Res Commun* **243**, 706-710 (1998).
29. Morgan, B.P., van den Berg, C.W., Davies, E.V., Hallett, M.B. & Horejsi, V. Cross-linking of CD59 and of other glycosyl phosphatidylinositol-anchored molecules on neutrophils triggers cell activation via tyrosine kinase. *Eur J Immunol* **23**, 2841-2850 (1993).
30. van den Berg, C.W., Cinek, T., Hallett, M.B., Horejsi, V. & Morgan, B.P. Exogenous glycosyl phosphatidylinositol-anchored CD59 associates with kinases in membrane clusters on U937 cells and becomes Ca(2+)-signaling competent. *J Cell Biol* **131**, 669-677 (1995).
31. Doupnik, C.A. et al. Time Resolved Kinetics of Direct G β 1 γ 2 Interactions with the Carboxyl Terminus of Kir3.4 Inward Rectifier K⁺ Channel Subunits. *Neuropharmacology* **35**, 923-931 (1996).
32. Svensson, H.G. et al. Contributions of Amino Acid Side Chains to the Kinetics and Thermodynamics of the Bivalent Binding of Protein L to Ig κ Light Chain. *Biochemistry* **43**, 2445-2457 (2004).

33. Zhang, Z. & Vachet, R.W. Kinetics of Protein Complex Dissociation Studied by Hydrogen/Deuterium Exchange and Mass Spectrometry. *Anal Chem* **87**, 11777-11783 (2015).
34. Liu, H. et al. Investigating interactions of the *Bacillus subtilis* spore coat proteins CotY and CotZ using single molecule force spectroscopy. *J Struct Biol* **192**, 14-20 (2015).
35. Gavutis, M., Jaks, E., Lamken, P. & Piehler, J. Determination of the Two-Dimensional Interaction Rate Constants of a Cytokine Receptor Complex. *Biophys J* **90**, 3345-3355 (2006).
36. van der Merwe, P.A., Brown, M.H., Davis, S.J. & Barclay, A.N. Affinity and kinetic analysis of the interaction of the cell adhesion molecules rat CD2 and CD48. *EMBO J* **12**, 4945-4954 (1993).
37. Sayre, P.H., Hussey, R. E., Chang, H. C., Ciardelli, T. L., Reinherz, E. L. Structural and binding analysis of a two domain extracellular CD2 molecule. *J Exp Med* **169**, 995-1009 (1989).
38. van der Merwe, P.A. et al. Human Cell-Adhesion Molecule CD2 Binds CD58 (LFA-3) with a Very Low Affinity and an Extremely Fast Dissociation Rate but Does Not Bind CD48 or CD59. *Biochemistry* **33**, 10149-10160 (1994).
39. Immordino, M.L., Dosio, F. & Cattel, L. Stealth liposomes: review of the basic science, rationale, and clinical applications, existing and potential. *Int J Nanomedicine* **1**, 297-315 (2006).

40. Du, H., Chandaroy, P. & Hui, S.W. Grafted poly-(ethylene glycol) on lipid surfaces inhibits protein adsorption and cell adhesion. *BBA Biomembranes* **1326**, 236-248 (1997).
41. Price, M.E., Cornelius, R.M. & Brash, J.L. Protein adsorption to polyethylene glycol modified liposomes from fibrinogen solution and from plasma. *BBA Biomembranes* **1512**, 191-205 (2001).
42. Torchilin, V.P. Recent advances with liposomes as pharmaceutical carriers. *Nat Rev Drug Discov* **4**, 145-160 (2005).
43. Criss, A.K. & Seifert, H.S. A bacterial siren song: intimate interactions between *Neisseria* and neutrophils. *Nat Rev Micro* **10**, 178-190 (2012).
44. Leusch, H.G., Drzeniek, Z., Markos-Pusztai, Z. & Wagener, C. Binding of *Escherichia coli* and *Salmonella* strains to members of the carcinoembryonic antigen family: differential binding inhibition by aromatic alpha-glycosides of mannose. *Infect Immun* **59**, 2051-7 (1991).
45. Virji, M. et al. Carcinoembryonic antigens are targeted by diverse strains of typable and non-typable *Haemophilus influenzae*. *Mol Microbiol* **36**, 784-795 (2000).
46. Hill, D.J. & Virji, M. A novel cell-binding mechanism of *Moraxella catarrhalis* ubiquitous surface protein UspA: specific targeting of the N-domain of carcinoembryonic antigen-related cell adhesion molecules by UspA1. *Mol Microbiol* **48**, 117-129 (2003).
47. Berger, C.N., Billker, O., Meyer, T.F., Servin, A.L. & 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* **52**, 963-983 (2004).
48. Barnich, N., Carvalho F. A., Glasser, A-L., Darcha, C., Jantsheff P., Allez M., Peeters H., Bommelaer, G., Desreumaux, P., Colombel, J-F., Darfeuille-Michaud, A. CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. *J Clin Invest* **117**, 1566-74 (2007).
 49. Blake, M.S., Blake, C.M., Apicella, M.A. & Mandrell, R.E. Gonococcal opacity: lectin-like interactions between Opa proteins and lipooligosaccharide. *Infect Immun* **63**, 1434-9 (1995).
 50. Fox, D.A. et al. Structure of the Neisserial Outer Membrane Protein Opa60: Loop Flexibility Essential to Receptor Recognition and Bacterial Engulfment. *JACS* **136**, 9938-9946 (2014).
 51. Minor, S.Y., and Gotschlich, E. C. The genetics of LPS synthesis by the gonococcus. in *Genetics of Bacterial Polysaccharides* (ed. Goldberg, J.B.) 111-131 (CRC Press, 1999).
 52. Whitesides, G.M. & Krishnamurthy, V.M. Designing ligands to bind proteins. *Q Rev Biophys* **38**, 385-395 (2005).
 53. Sethi, A., Goldstein, B. & Gnanakaran, S. Quantifying Intramolecular Binding in Multivalent Interactions: A Structure-Based Synergistic Study on Grb2-Sos1 Complex. *PLoS Comput Biol* **7**, e1002192 (2011).
 54. Mammen, M., Choi, S.-K. & Whitesides, G.M. Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew Chem Int Ed* **37**, 2754-2794 (1998).

55. Crothers, D.M. & Metzger, H. The influence of polyvalency on the binding properties of antibodies. *Immunochemistry* **9**, 341-357 (1972).
56. Huskens, J. Multivalent interactions at interfaces. *Curr Opin Chem Biol* **10**, 537-543 (2006).
57. Ercolani, G., Piguet, C., Borkovec, M. & Hamacek, J. Symmetry Numbers and Statistical Factors in Self-Assembly and Multivalency. *J Phys Chem B* **111**, 12195-12203 (2007).

5. Prospects for future research into the molecular determinants of Opa – CEACAM interactions

5.1 Overview

This dissertation presented the affinities of many Opa protein – CEACAM receptor interactions *in vitro*. A tight affinity was found for all Opa – CEACAM interactions investigated. While knowledge of the affinities of Opa – CEACAM interactions help us to better understand how *Neisseria* adhere to and gain entry into host cells, many questions about this interaction still remain. Does receptor affinity correlate with *Neisseria* invasion efficacy? How do diverse sequences of Opa proteins engage the same receptor? What Opa residues are involved in receptor interactions? We seek to answer these questions to gain insight into the mechanism of molecular recognition and the specific molecular determinants of Opa – CEACAM interactions. Additionally, we seek to determine a high-resolution structure model of the Opa-CEACAM complex. Progress of ongoing research towards this goal is presented, as well as future prospects.

5.2 Mapping the regions of Opa HV1 and HV2 that interact with CEACAM

Previous investigations determined which residues of CEACAM1 are involved in interactions with various Opa proteins.¹⁻⁶ The surface area of the amino acid residues involved in interactions with Opa is $\sim 400 \text{ \AA}^2$. Therefore, we propose only a small number of Opa residues, approximately five total, on both HV1 and HV2 regions are directly involved in interactions with CEACAM. We hypothesize that several conserved hydrophobic residues in HV1 and HV2, as well as proline residues in HV2, are essential in engaging CEACAM receptors, either by stabilizing HV1 – HV2 interactions to form a binding pocket for the

CEACAM, and/or directly interacting with CEACAMs. To identify this small number of amino acids from among the ~ 65 that compose both the HV1 and HV2 region, we will utilize several techniques including competition experiments via fluorescence polarization, and protein footprinting via limited proteolysis protection.

5.2.1 Competition assays using FP to determine which regions on Opa HV loops are involved in receptor interactions

We postulate that only a few amino acids on extracellular loops two and three of Opa proteins are involved in receptor recognition and binding. Loops two and three of the Opa family of proteins contain regions of hypervariable sequences, and are referred to as HV1 and HV2, respectively. To investigate which specific residues are important in the Opa-CEACAM interaction, binding competition studies are being conducted using synthetic peptides with sequences that mimic the variable regions in the HV1 and HV2 loops of the Opa₆₀ protein (Figure 5.1). Thus far, one HV1 peptide (termed HV1_1) and two HV2 peptides (termed HV2_1 and HV2_2) were tested for competition of the CEACAM binding domain.

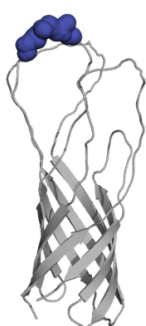
Fluorescence polarization experiments were carried out as described previously, using a constant fl-NCCM1 concentration of 5 nM and a constant Opa₆₀ concentration of 50 nM. Peptide concentration was varied for optimization. The fraction of NCCM1 bound to Opa proteoliposomes was decreased by the addition of the HV peptides: peptides HV1_1 and HV2_2 both decreased the fraction of CEACAM bound by ~10%, while the HV2_1 peptide alone decreased binding by ~20% (Figure 5.2). A combination of HV1_1 peptide with HV2_2 peptide demonstrated a decrease in binding of ~35%, while HV1_1 combined with HV2_1 did not demonstrate a further decrease in binding (Figure 5.2).

The largest impact on CEACAM – Opa binding was seen with the addition of both the HV1_1 and HV2_2 peptide in the reaction mixture, indicating that both HV1 and HV2 loop regions are involved in interacting with CEACAMs. Both the HV1_1 peptide and the HV2_2 peptide alone decreased the Opa – CEACAM binding by ~ 10%, so if the affect was additive we would expect to see a decrease of ~ 20% with the addition of both peptides. However, using HV1_1 and HV2_2 combined, the decrease in binding more than doubled, which may indicate that interactions between HV1_1 and HV2_2 have a larger impact on Opa – CEACAM binding. Bos *et al* had previously shown that a unique combination of both HV1 and HV2 loop sequences are necessary for CEACAM interactions, which our data further supports.⁷ Importantly, this data does not imply that these Opa HV1 and HV2 residues are directly interacting with NCCM1. Rather, we can only assert that residues in these regions have an effect on the Opa – CEACAM interaction, which may be due to indirect modulations of Opa loop conformations.

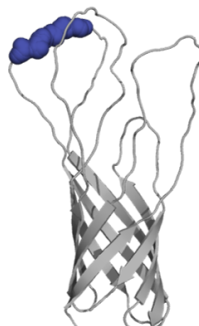
After more Opa₆₀ HV loop peptides are synthesized and investigated for competitive binding, we aim to collect FP binding data using HV peptides containing single amino acid mutations to determine which specific residues are important for receptor interaction. However, there are some limitations to using this technique. Solubility of the Opa₆₀ HV peptides is a major concern, as these regions contain a high amount of hydrophobic amino acid residues, which we believe are involved in interacting with CEACAM. Additionally, as the fl-NCCM1 is already quite large for the limit of FP, it is difficult to distinguish if the NCCM1 is bound to the Opa proteoliposomes or if the NCCM1 is interacting with both HV1 and HV2 peptides in our competition experiments. As such, we are employing additional methods to

elucidate the specific amino acid residues on the Opa HV regions directly involved in CEACAM binding.

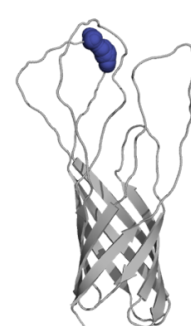
Opa ₆₀ peptide sequences for competitive binding studies	
Peptide name	Opa ₆₀ HV1 sequences
	NNKYSVNIENVRIKENGIRIDRK
HV1_1	ENVRIKENG
Peptide name	Opa ₆₀ HV2 sequences
	SIDSTKKTIEVTTVPSNAPNGAVTTYNTDPKTQNDYQSNSI
HV2_1	EVTTVPSNA
HV2_2	NGAVTTYNTD



HV1_1



HV2_1



HV2_2

Figure 5.1. Sequences of the Opa₆₀ HV regions used for competition experiments. Three peptides were chosen based on the Opa₆₀ HV 1 and 2 regions, termed HV1, HV2_1, and HV2_2 (top). The Opa₆₀ structure is shown (grey) with the amino acids corresponding to the peptides highlighted in blue (bottom).

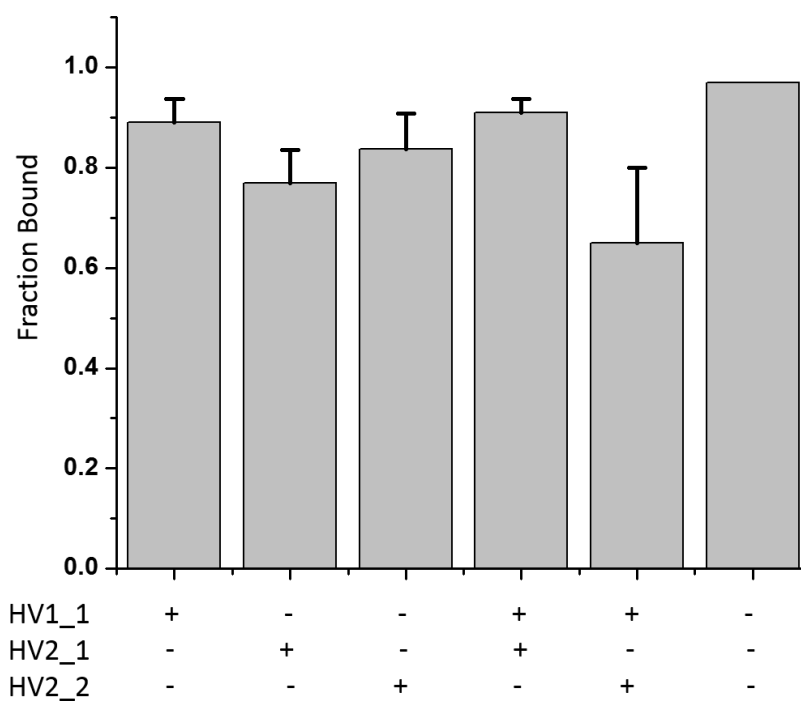


Figure 5.2. Peptide competition experiments indicate decreased Opa – CEACAM interactions with peptide present. Opa – NCCM binding was assessed in the presence (+) or absence (-) of HV peptides. Data is graphed as the fraction of Opa bound to the fl-NCCM1, normalized to the Opa₆₀ – NCCM1 fully bound state (absence of all HV peptides). FP conditions contained 5 nM fl-NCCM1, 50 nM Opa₆₀ (based off of FP data shown in Chapter 3), with a total peptide concentration of 125 μ M.

5.2.2 Protein footprinting to assess which amino acid residues on the loops of Opa proteins are involved in binding to receptors

In order to assess which residues on the Opa HV regions are involved in receptor interactions, we are utilizing limited proteolysis to compare cleavage of Opa₆₀ liposomes with and without NCCM1. By incubating Opa₆₀ proteoliposomes with NCCM1 before the addition of a protease, amino acid residues in the Opa HV1 and HV2 regions that would normally be cleaved by the protease can be protected by the bound NCCM1. Therefore, we would see limited or no cleavage around those residues on Opa near the CEACAM-binding interface. We are using SDS-PAGE as an initial screen for protease protection, and then will utilize mass spectrometry to analyze the peptide fragments created from the cleavage.

Opa₆₀ proteoliposomes were treated with trypsin in the presence and absence of NCCM1 (Figure 5.3), and assessed for the appearance of the Opa barrel cleavage product (The Opa β -barrel is resistance to SDS denaturation, even when the loops are removed via trypsin,⁸ see Chapter 2). At each time point there is less cleaved product in the presence of NCCM1 than in the sample without NCCM1, indicating that regions of the Opa extracellular loops are being protected from trypsin cleavage. Predicted cleavage patterns of Opa loops two and three show that we can get relatively complete coverage of both HV1 and HV2 regions with only a few different proteases (Figure 5.4).

After proteolysis, the generated Opa₆₀ fragments will be analyzed using mass spectrometry (MS) to detect differences in the cleavage pattern caused by the addition of NCCM1. Figure 5.5 provides a sample of the type of data that would be expected, based on our hypothesis about the hydrophobic residues in the HV regions of Opa₆₀. For example, if Y82, V84, V89, I91, I148, and V157 are involved in CEACAM interactions, then there would

be a decrease in the cleavage around those specific residues (Figure 5.5). Although not necessary, NCCM1 is largely resistant to proteolytic cleavage, which will aid in the analysis of the proteolytic fragments.

The resolution of this experiment may be limited, but Lacy *et al* were able to detect differences in cleavage locations one residue apart using the same techniques.⁹ However, if we are only able to resolve smaller regions of HV1 and HV2 instead of specific residues involved in interactions with CEACAMs, additional protection assays (such as chemical modification, described by Giovanni *et al*)¹⁰, will be undertaken as well using a similar approach.

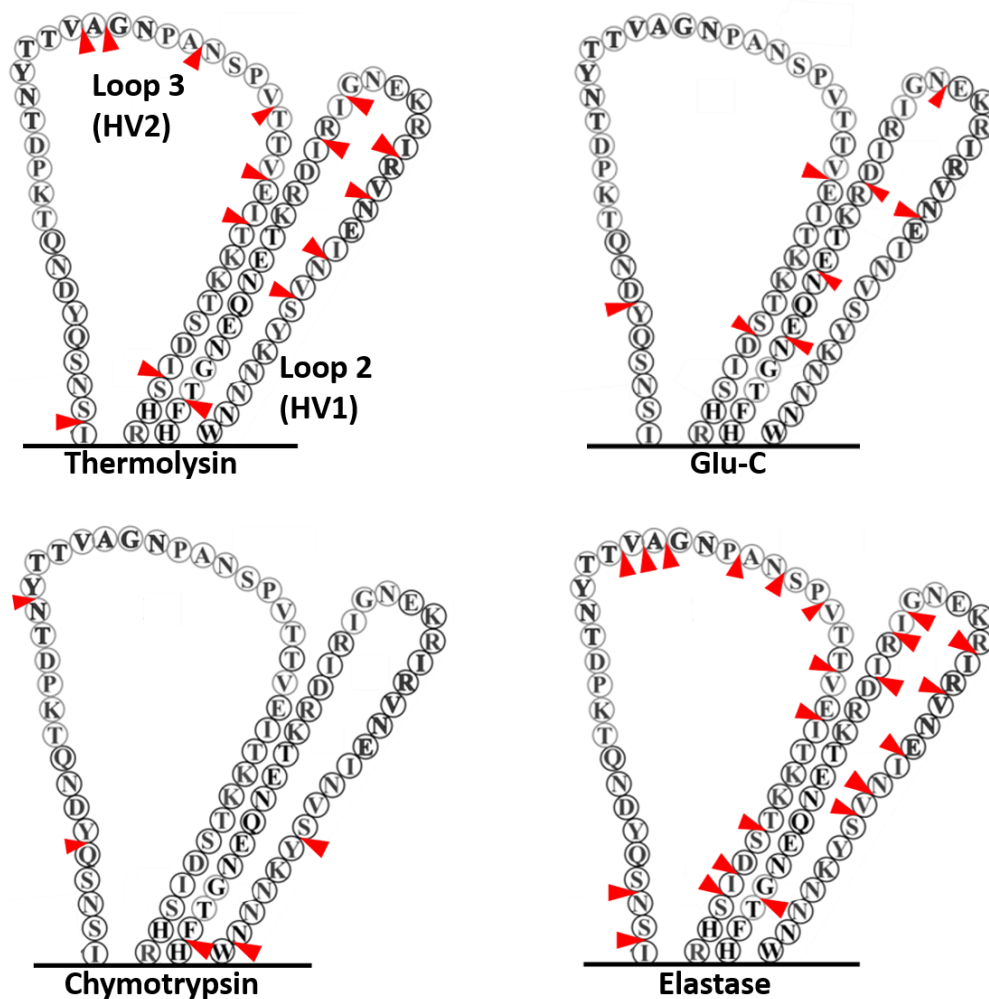


Figure 5.4. Cleavage patterns of Opa₆₀ loops. A variety of proteases will be utilized to cleave Opa₆₀ with and without the addition of NCCM1 to get complete coverage of both HV regions. Predicted cleavage sites are indicated with red arrows.

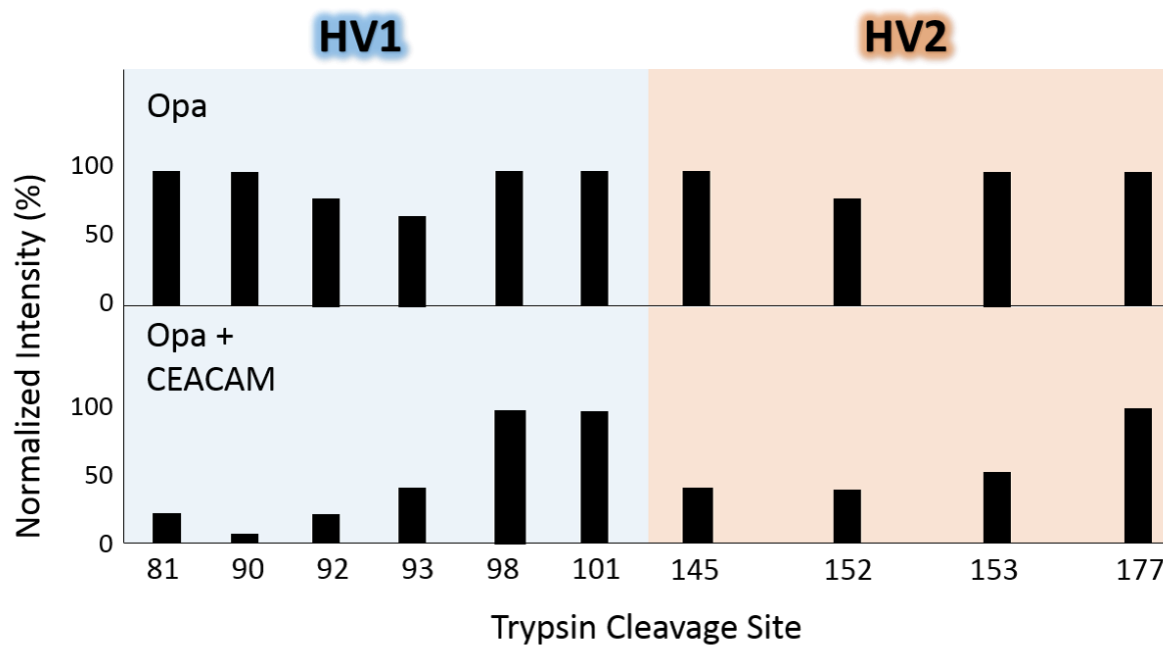


Figure 5.5. Theoretical mass spectrometry data for trypsin cleavage of Opa with and without NCCM. Cleavage near the Opa – CEACAM binding site will not be observed because CEACAM is interacting with amino acid residues in that region, occluding protease access. Numbers correspond to amino acid residues on Opa₆₀ where trypsin would normally cleave.

5.3 Determining the structure of the Opa – CEACAM complex

5.3.1 Progress towards structure determination

We seek to determine a high-resolution structure model of the Opa-CEACAM complex in detergent to gain a clearer understanding of the interactions between these two proteins. Towards this end, sparse crystal screen trays have been set, which contain both Opa₆₀ folded into detergent and NCCM1, to investigate various crystallization buffer conditions. Trays are set to be kept at different temperatures, with different protein concentrations, with Opa folded into a variety of detergents that have been previously successful with membrane protein crystallizations (*N,N*-dimethyldodecylamine-*N*-oxide (LDAO), *n*-octyl- β -d-glucopyranoside (OG), octyltetraoxyethylene (C₈E₄), *n*-decyl- β -d-maltopyranoside (DM) and *n*-dodecyl- β -d-maltopyranoside (DDM)).^{11,12} Several of the conditions have given promising crystalline hits (examples in Figure 5.6). Finer screen trays have been set for several of the conditions that provided the best crystalline leads.

All developed crystals (approximately ten) were taken to the APS 24-ID-C beamline by Kimberly Stanek of the Mura Laboratory. The majority of crystals contained only salt, or dissolved before data could be collected. However, a full data set was collected to 3.25 Å resolution on the crystal shown in the middle panel of Figure 5.6, with the crystal buffer conditions of 0.2 M Ca(C₂H₃O₂)₂, 20 % w/v PEG3350, with Opa folded in OG, at 4 °C. The collected data set was processed by Peter Randolph, also of the Mura Laboratory, in the P6₄22 space group. Unfortunately, only NCCM1 was present in the electron density (Figure 5.7, A). In this crystal, NCCM1 packed as a tetramer. Amino acids involved in binding to Opa (ex. Y68 and I125)¹³ are occluded in this structure (Figure 5.7, B-D), ruling out the possibility of soaking the crystal with Opa HV peptides. The previously determined NCCM1 structure was

processed in the P6₃ space group as a dimer, with 2.2 Å resolution, however, the amino acids on the Opa-binding face are buried in this structure as well.⁶

5.3.2 Prospects for future research

Co-crystallization of a membrane protein – receptor complex is extremely challenging, with potential difficulties ranging from solubilization to solving phases. Detergent screening and complex formation are the two most important factors at this stage. Alternative approaches to crystallizing the whole complex would be to co-crystallize NCCM1 with peptides identified from the FP peptide competition experiments or by protein foot printing, with co-crystallization screens performed at different ratios of NCCM1 to peptide.

Crystallizing the NCCM1 domain in a conformation with the Opa-binding face exposed is an additional approach that we are undertaking. We have set several GST-NCCM1 trays to increase the likelihood of obtaining crystals with the Opa-binding residues surface exposed. Once crystals are obtained, peptide soaks will be performed, using Opa HV peptides based off of results from the FP competition and protein footprinting experiments. Alternate lipid environments for the folded Opa proteins, bicelles and nanodiscs in particular, will also be investigated for use as an alternative to the detergent micelles.

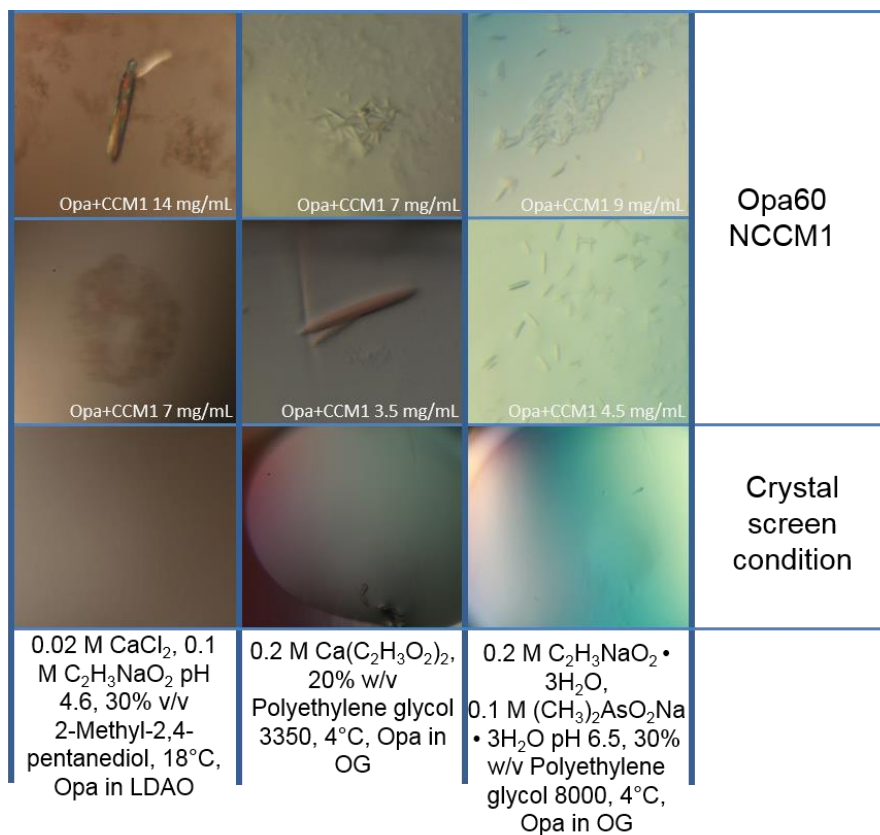


Figure 5.6. A selection of preliminary crystallization leads. Data are representative of several crystalline hits. Crystallization conditions are varied, including the crystallization buffer, temperature, and Opa – CEACAM concentration.

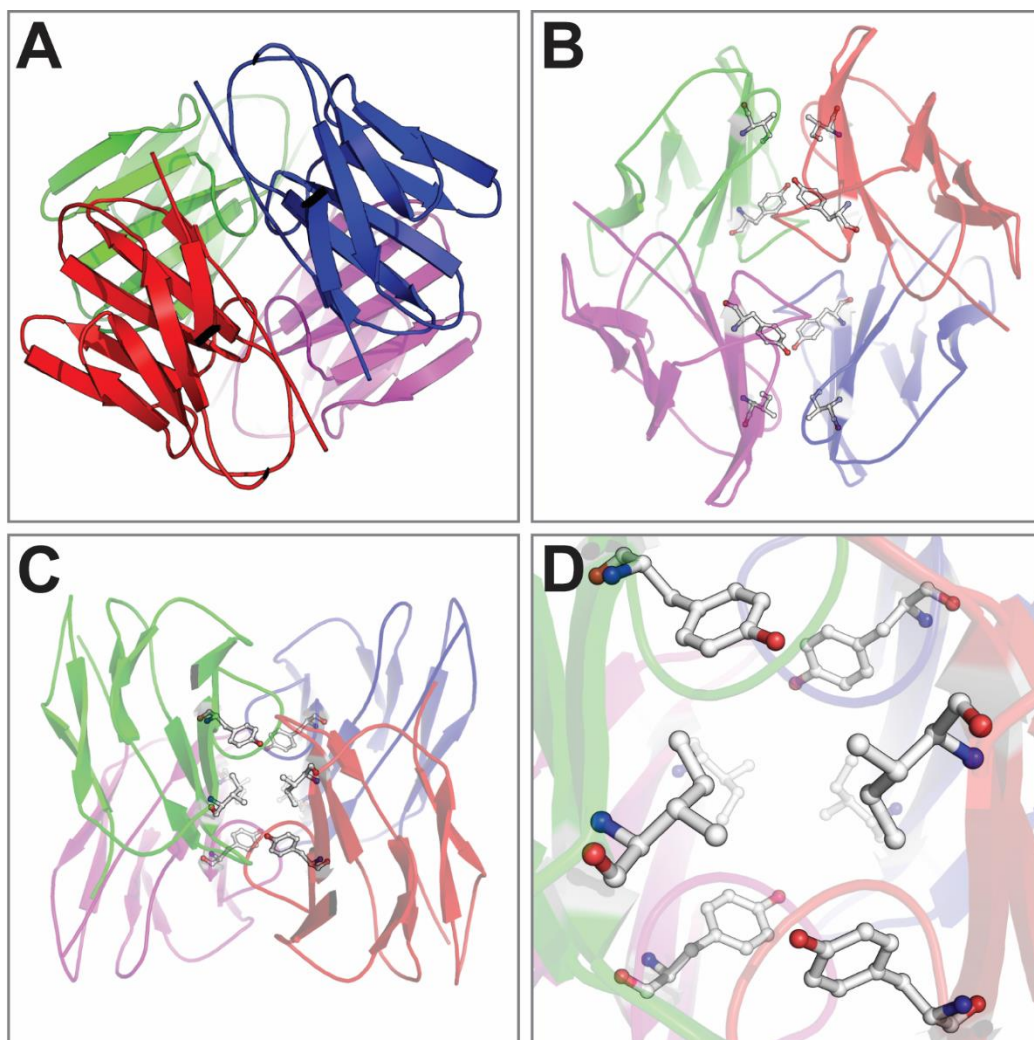


Figure 5.7. Crystal packing of NCCM1. The NCCM1 asymmetric unit is a tetramer (A-C), with the Opa binding faces of each monomer occluded in the center of the tetramer (B-D). NCCM1 is colored by subunit; Y68 and I125 are rendered as sticks (B-D). Image courtesy of Peter Randolph.

5.4 Long term applications of this research

Liposomes have been clinically established as therapeutic delivery systems. To date, liposomes have been utilized for the delivery of antifungals, antimicrobials, genes, and vaccines.^{14,15} Liposomes are particularly well suited for use as delivery vehicles, as cargo can be encapsulated in the aqueous core, or embedded into the hydrophobic region of the bilayer. However, major problems with liposomal delivery systems on the market include targeting specific cells, and delivering cargo intracellularly.^{16,17}

As certain bacterial membrane proteins function to cause bacterial engulfment by the host cell, these proteins may be exploited to create novel therapeutic delivery systems that can overcome many of the challenges listed above. Specifically, we are interested in designing functionalized liposomes based on the Opa binding mechanism, as the interaction of Opa with receptors is specific and triggers engulfment of the cell. Using recombinant Opa proteoliposomes as a therapeutic delivery system is not practical, as the stability is low and the cost would be high. However, when amino acid residues on the HV regions of Opa involved in interactions with CEACAMs are elucidated, we can design peptide decorated liposomes to mimic the effect.

5.5 Concluding remarks

Bacterial resistance is a global threat to public health. In many cases, specific interactions between membrane proteins on the bacteria and human host cells lead to bacterial colonization and engulfment. Knowledge of these interactions will advance our understanding of bacterial pathogenesis, as well as provide the opportunity to use the bacterial mechanisms to our own advantage, such as in developing vaccines and new methods of treatment.

While a number of Opa variants have been classified in terms of receptor specificity *in vivo*, a receptor recognition motif has yet to be identified. Multiple sequence alignments of the Opa HV regions do not provide any insight into receptor recognition. Elucidating the molecular determinants of the Opa protein – CEACAM receptor interactions will provide the foundation for the design of a targeted liposome therapeutic delivery system for human cells.

Binding affinities have been determined for a variety of Opa proteins and CEACAM receptors. Phagocytosis of the Opa proteoliposomes is being investigated to determine if there is a correlation between receptor specificity and bacterial engulfment. Additionally, we seek to determine a structure of the Opa protein – CEACAM receptor complex, which will provide insight into Neisserial pathogenicity.

5.6 References

1. Virji, M. et al. Critical determinants of host receptor targeting by *Neisseria meningitidis* and *Neisseria gonorrhoeae*: identification of Opa adhesiotopes on the N-domain of CD66 molecules. *Mol Microbiol* **34**, 538-551 (1999).
2. Popp, A., Dehio, C., Grunert, F., Meyer, T.F. & Gray-Owen, S.D. Molecular analysis of neisserial Opa protein interactions with the CEA family of receptors: identification of determinants contributing to the differential specificities of binding. *Cell Microbiol* **1**, 169-181 (1999).
3. Virji, M., Watt, S.M., Barker, S., Makepeace, K. & Doyonnas, R. The N-domain of the human CD66a adhesion molecule is a target for Opa proteins of *Neisseria meningitidis* and *Neisseria gonorrhoeae*. *Mol Microbiol* **22**, 929-39 (1996).
4. Bos, M.P., Hogan, D. & Belland, R.J. Homologue scanning mutagenesis reveals CD66 receptor residues required for neisserial Opa protein binding. *J Exp Med* **190**, 331-40 (1999).
5. Bos, M.P., Kuroki, M., Krop-Watorek, A., Hogan, D. & Belland, R.J. CD66 receptor specificity exhibited by neisserial Opa variants is controlled by protein determinants in CD66 N-domains. *Proc Natl Acad Sci U S A* **95**, 9584-9589 (1998).
6. Fedarovich, A., J. Tomberg, R. A. Nicholas, and C. Davies. Structure of the N-terminal domain of human CEACAM1: binding target of the opacity proteins during invasion of *Neisseria meningitidis* and *N. gonorrhoeae*. *Acta Crystallogr Sec D-Biol Crystallogr* **62**, 971-979 (2006).

7. Bos, M.P., Kao, D., Hogan, D.M., Grant, C.C. & Belland, R.J. Carcinoembryonic antigen family receptor recognition by gonococcal Opa proteins requires distinct combinations of hypervariable Opa protein domains. *Infect Immun* **70**, 1715-23 (2002).
8. Fox, D.A. & Columbus, L. Solution NMR resonance assignment strategies for β -barrel membrane proteins. *Protein Sci* **22**, 1133-1140 (2013).
9. Lacy, E.R. et al. Molecular Basis for the Specificity of p27 Toward Cyclin-dependent Kinases that Regulate Cell Division. *J Mol Biol* **349**, 764-773 (2005).
10. Giovanni, R., Anna Maria, S., Simona, A., Chiara, D.A. & Andrea, S. Mass Spectrometry-Based Approaches for Structural Studies on Protein Complexes at Low-Resolution. *Curr Proteomics* **4**, 1-16 (2007).
11. Privé, G.G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397 (2007).
12. Newby, Z.E. et al. A general protocol for the crystallization of membrane proteins for X-ray structural investigation. *Nat Protoc* **4**, 619-37 (2009).
13. Virji, M. et al. Critical determinants of host receptor targeting by *Neisseria meningitidis* and *Neisseria gonorrhoeae*: identification of Opa adhesiotopes on the N-domain of CD66 molecules. *Mol Microbiol* **34**, 538-51 (1999).
14. Al-Jamal, W.T. & Kostarelos, K. Liposomes: From a Clinically Established Drug Delivery System to a Nanoparticle Platform for Theranostic Nanomedicine. *Acc Chem Res* **44**, 1094-1104 (2011).
15. Torchilin, V.P. Recent advances with liposomes as pharmaceutical carriers. *Nat Rev Drug Discov* **4**, 145-160 (2005).

16. Torchilin, V.P., Rammohan, R., Weissig, V. & Levchenko, T.S. TAT peptide on the surface of liposomes affords their efficient intracellular delivery even at low temperature and in the presence of metabolic inhibitors. *Proc Natl Acad Sci U S A* **98**, 8786-91 (2001).
17. Dutta, R.C. Drug carriers in pharmaceutical design: promises and progress. *Curr Pharm Des* **13**, 761-9 (2007).

Appendix: Materials and methods

Expression and purification of the N-terminal domain of CEACAMs 1 and 3.

The procedure for CEACAM expression and purification was adapted from Fedarovich *et al*¹. *E. coli* MC1061 cells transformed with a modified pGEX-2T plasmid (pGEX-2V) containing the N-terminal D1 domain of human *ceacam1* gene (amino acids 35-141 (MW = 11.8 kDa) of the mature protein and referred to as NCCM1, Figure S2) were generously provided by Rob Nicholas (University of North Carolina, Chapel Hill, NC.). The N-terminal domain of human *ceacam3* gene (amino acids 35-142 (MW = 12.2 kDa) of the mature protein and referred to as NCCM3, Figure S2) was synthesized and cloned into a pGEX-2T vector (Bio Basic Inc., Ontario, Canada). To maintain a construct similar to that of NCCM1, a linker region between the GST and NCCM3 was designed to incorporate a tobacco etch virus protease (TEV) cleavage site (ENLYFQ | G) in the resulting fusion protein. Additionally, the amino acids SGA were added as a spacer immediate following the TEV cleavage site and before the first amino acid of NCCM3. NCCM cysteine mutations were introduced using PIPE mutagenesis.²

MC1061 *E. coli* cells with CEACAM plasmid were grown in Luria-Bertani (LB) media supplemented with streptomycin and ampicillin (50 µg/mL each) at 37°C until an OD₆₀₀ ≈ 0.6 was reached. Cell cultures were cooled to 25°C and protein expression was induced with 1 mM isopropyl β-thio-D-galactoside (IPTG) overnight with constant shaking (200 rpm) at the same temperature. Cells from 1 L culture were harvested by centrifugation (4,500 x g, 20 min, 4°C), resuspended in 15 mL lysis buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 2 mM ethylenediaminetetraacetic acid (EDTA), 2 mM dithiothreitol (DTT), 10% glycerol, half of a Complete protease inhibitor tablet (Roche)), and lysed using a microfluidizer (Microfluidics model 110L, Newton, MA). Cell debris was removed by centrifugation (18,000 x g, 1 h, 4°C),

and proteins were precipitated from the supernatant by the addition of ammonium sulfate to 55% saturation with constant stirring for 1 h at 4°C. Precipitated proteins were harvested by centrifugation (12,000 x g, 30 min, 4°C), pellets were resuspended in 30 mL lysis buffer, and added to a glutathione resin column previously equilibrated with equilibration buffer (20 mM Tris, pH 7.3, 150 mM NaCl, 2 mM DTT, and 10% glycerol) at 4°C. After loading, the column was washed with 10 column volumes of equilibration buffer, and eluted with 50 mL of the same buffer supplemented with 10 mM reduced glutathione. To cleave the N-terminal domain from the GST, TEV (~ 3.5 μ M) was added to the eluent containing purified GST-NCCM fusion protein and dialyzed (20 mM Tris, pH 7.3, 150 mM NaCl, 2mM DTT, 10% glycerol; MWCO = 3,500 kDa) overnight at 4°C. The N-terminal domain of CEACAM was purified from GST and TEV by using an HR Sephacryl S-200 gel-filtration column (26/60 mm, GE Healthcare) previously equilibrated with 20 mM Tris, pH 8.0, 500 mM NaCl, 2 mM DTT, and 10% glycerol. Fractions containing pure NCCM (as assessed by SDS-PAGE) were combined and concentrated to ~42 μ M (determined by A_{280} ; ϵ = 14,440 $M^{-1}cm^{-1}$ for NCCM1 and ϵ = 15,930 $M^{-1}cm^{-1}$ for NCCM3) and stored at -80°C. Wild-type NCCMs formed SDS resistant artificial dimers before experiments could be carried out, so in all cases a Cysteine mutant was used (H139C), which helped to slow the dimerization process. Protein purity was greater than 95%, as assessed by SDS-PAGE.

Labeling NCCM with fluorescent probe.

Purified NCCM1 or NCCM3 H139C was dialyzed overnight at 4°C to remove DTT (20 mM Tris, pH 7.3, 150 mM NaCl, 10% glycerol; MWCO = 3,500 kDa), and concentrated to ~50 μ M. A 1 mM stock of the fluorescent dye, 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic

acid, disodium salt (AMS, Life Technologies, Carlsbad, CA), was freshly prepared and added drop wise to the protein to yield a dye: protein molar ratio of 20:1. The reaction was protected from light and carried out under nitrogen overnight at 4°C. Excess dye was removed by extensive dialysis (MWCO = 3,500 kDa) at 4°C. Labeling efficiency was determined to be greater than 95%, as assessed by MALDI-TOF mass spectrometry.

Cloning and expression of N. gonorrhoeae.

FA1090 Opa⁻ and FA1090 OpaD_{nv} Gc, which is phase-locked ON for constitutive OpaD expression (OpaD⁺), were generated as previously described.³ FA1090 Opa50_{nv} and FA1090 Opa60_{nv} were constructed by transformation of FA1090 Opa⁻ with a plasmid (pST) containing the non-variable signal sequence of OpaD (99bp) immediately followed by sequences corresponding to the mature Opa50 (711bp) or Opa60 (717bp) protein from MS11 Gc, and flanked by 730bp of the genomic sequence 5' of OpaD and 889bp 3' of OpaD from Opa⁻ Gc. These constructs were synthesized and ligated into the pST vector by Genewiz Inc (South Plainfield, NJ). Successful transformants were selected for by opaque colony morphology and confirmed by PCR, sequencing and immunoblot.

Cloning, expression and purification of recombinant Opa proteins for liposomes.

The *opa60*, *opa50*, and *opaD* genes were sub-cloned into pET28b vectors (EMD chemicals, Gibbstown, NJ) encoding a thrombin cleavable N-terminal His₆-tag (MGSSHHHHHSSGLVPRGSHM). Expression and purification protocols were followed as previously described.⁴⁻⁶ Briefly, the *opa* containing plasmids were transformed into a BL21(DE3) *E. coli* strain. Cell cultures were grown in LB media and expression was induced to

the insoluble fraction. Cells were harvested, resuspended in lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl, half of a Complete protease inhibitor pellet), and lysed. The insoluble fraction was pelleted. The pellet was resuspended in extraction buffer (lysis buffer with 8 M urea) and the insoluble fraction was removed by centrifugation. Opa proteins were purified using Co^{2+} immobilized metal affinity chromatography, and eluted (20 mM sodium phosphate, pH 7.0, 150 mM NaCl, 680 mM imidazole, and 8 M urea). The eluted fractions containing Opa were pooled and concentrated (MWCO = 10 kDa) to 0.5 mM, with an expression yield of ~20 mg/L of cell culture. Protein concentration was determined by A_{280} ($\epsilon = 41,830 \text{ M}^{-1}\text{cm}^{-1}$ for Opa₆₀, $\epsilon = 43,320 \text{ M}^{-1}\text{cm}^{-1}$ for OpaD, $\epsilon = 40,340 \text{ M}^{-1}\text{cm}^{-1}$ for Opa₅₀) and purity was greater than 95%, assessed by SDS-PAGE.

N. gonorrhoeae pull-down assays.

Gc were grown on gonococcal medium base (GCB, Difco, Becton-Dickinson, Franklin Lakes, NJ) containing Kellogg's supplements I and II (Kellogg, 1963) for approximately 8 hours, then grown in rich liquid medium (GCBL) with periodic dilutions to produce uniformly mid-logarithmic cultures, as previously described.⁷ Using OD_{550} to calculate, approximately 3×10^8 CFU/mL Gc was suspended in 1-2 mL GCBL. Cultures were spiked with either NCCM1 (3 mg/mL) or GST-NCCM3 (28 mg/mL) and incubated at 37°C for 30 min with rotation. Cultures were centrifuged (500 x g for 20 min) and 800 μL of supernatant added to 200 μL 5X SDS-PAGE loading buffer. Pellets were washed twice by resuspending in 1 mL PBS + 5mM MgSO_4 and centrifuging at 10,000 x g for 3 min. Pellets were resuspended in 200 μL 1X SDS-PAGE loading buffer. Pellet lysates and supernatant samples were analyzed by SDS-PAGE and immunoblotting.

Liposome pull-down assays.

Opa protein folding and reconstitution was adapted from Dewald *et al.*⁵ Lipid stocks of 1,2-didecanoyl-*sn*-glycero-3-phosphocholine (diC₁₀PC, Avanti Polar Lipids, Alabaster, AL), originally dissolved in chloroform, were dried under a continuous stream of nitrogen and resuspended into borate buffer (10 mM sodium borate, pH 12, 1 mM EDTA) and sonicated for 30 minutes (Q Sonica model Q500, Newtown, CT) with a 1/8 inch micro tip at 40% amplitude. Post-sonication, urea was added to a concentration of 4 M, and purified recombinant unfolded Opa protein was added in 20 μ L aliquots, mixing between additions, yielding a final protein to lipid ratio of 1:1160. The folding reaction was incubated at 37°C for 3 days after which folding was assessed by SDS-PAGE. After folding, Opa proteoliposomes were harvested by ultracentrifugation (142,400 x g, 2 h, 10°C), resuspended in a new lipid mixture in resuspension buffer (30 mM Tris, pH 7.3, 150 mM NaCl), and pulse sonicated (30 s on, 30 s off for a total of 20 min). The final lipid composition contained 63 mol% 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC), 16 mol% 1,2-dimyristoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol) (sodium salt) (DMPG), 16 mol% cholesterol, 5 mol% 1,2-dimyristoyl-*sn*-glycero-3-phosphoethanolamine-N-[methoxy(polyethylene glycol)-1000] (ammonium salt) (DMPE PEG 1000), with a protein to lipid ratio of 1:234. Opa protein concentration was determined via A₂₈₀, blanking against empty liposomes, and 3.4 μ mol of the Opa proteoliposome solutions were aliquoted. Opa_{Trp} liposomes were prepared by adding trypsin to the Opa-liposome solution, 10 μ g of trypsin for every 2 μ g of Opa. The trypsin-liposome solution was incubated at room temperature for ~ 4 h. After incubation, the trypsin – liposome solution was passed over a column containing p-Aminobenzamidine-agarose (Sigma, St. Louis, MO) previously equilibrated with Opa resuspension buffer to remove trypsin. The flow through was collected and

cleavage was assessed by SDS-PAGE (Figure S4). CEACAM was added in excess (10.2 μ mol) and samples were incubated for 30 min at room temperature. Proteoliposomes were harvested by ultracentrifugation (142,400 \times g, 2.5 h, 10°C), and resuspended in 1 mL PBS. Ninety μ L of sample was added to 30 μ L of 4x SDS loading buffer, and analyzed via SDS-PAGE and immunoblotting.

SDS-PAGE and Immunoblotting.

All samples were boiled for 20 min and homogenized by passing through a needle (30 gauge) and syringe ten times. Twenty μ L were loaded onto a pre-cast 4-20% acrylamide gel (Bio-Rad, Hercules, CA). After electrophoresis, gels were transferred to 0.2 μ M nitrocellulose using the turbo blot system (Bio-Rad), and immunoblotted with the A0115 α -CEACAM antibody (Dako, Carpinteria, CA), or the 4B12 pan-Opa antibody (hybridoma generously provided by Christof Hauck, University of Konstanz, Germany, and antibody purified by the University of Virginia Hybridoma Core) followed by IRDye 800 CW conjugated goat anti-mouse or goat anti-rabbit IgG (LI-COR Biosciences, Lincoln, NE). Blots were visualized with an Odyssey imager (LI-COR Biosciences).

Fluorescence polarization binding assays.

Fluorescently labeled NCCM1 or NCCM3 (5 nM) was added to 300 nM Opa protein, and serially diluted with buffer (5 nM NCCM, 20 mM Tris, pH 7.3, 150 mM NaCl) across 11 steps, and done in triplicate. Samples were incubated in an opaque 96-well plate for 30 min, and then the fluorescence polarization was measured for each sample. Fluorescence polarization spectra were collected with a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA). The

excitation wavelength was 323 nm and emission spectra were collected at 411 nm, with a wavelength cutoff of 325 nm. Polarization was measured and converted into the fraction of CEACAM bound to the proteoliposomes at varying concentrations of Opa protein. Data was processed and analyzed using Origin Pro 7.5, using the following equation:

$$f_B = \left(\frac{K_d + R_T + L_T - \sqrt{(K_d + R_T + L_T)^2 - 4R_T L_T}}{2L_T} \right) \quad \text{eq.1}$$

Where f_B is the fraction bound, K_d is the ligand dissociation constant, R_T is the receptor concentration, and L_T is the total fluorescent ligand concentration. This equation takes ligand depletion into account, as described by Veiksina *et al.*⁸

Biolayer interferometry binding assays

BLI binding assays were carried out on an Octet Red96 system (fortéBIO, Menlo Park, CA). The GST fusion construct of CEACAMa 1, 3, 4, 5, and 8 were purified as described above, and incubated for 180s with Anti-GST biosensors (fortéBIO) that had been previously equilibrated with Opa resuspension buffer. The biosensors were washed in Opa resuspension buffer for 60 s. The biosensors were then incubated with 200 nM Opa proteoliposomes for 180 s. Biosensors were then dipped in Opa resuspension for 600 s. Raw data was background subtracted using empty liposomes as a reference. Data was processed using Savitzky-Golay filtering to remove high frequency noise from the data, and analyzed using ForteBio Data Analysis Software version 8.2 with a 1:1 model

FP competition experiments using HV peptides

FP competition experiments were carried out as described above. HV peptides were synthesized by Genscript (Piscataway, NJ). [Fl-NCCM] was constant at 5 nM, [Opa₆₀] was held constant at 50 nM, and 125 μ M total peptide concentration was added to the mixture. The fraction of CEACAM bound was based off of a reference samples at saturating conditions.

References

1. Fedarovich, A., J. Tomberg, R. A. Nicholas, and C. Davies. Structure of the N-terminal domain of human CEACAM1: binding target of the opacity proteins during invasion of *Neisseria meningitidis* and *N. gonorrhoeae*. *Acta Crystallogr Sect D-Biologic Crystallogr* **62**, 971-979 (2006).
2. Klock, H.E. & Lesley, S.A. The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. *Methods Mol Biol* **498**, 91-103 (2009).
3. Ball, L.M. & Criss, A.K. Constitutively Opa-expressing and Opa-deficient *Neisseria gonorrhoeae* strains differentially stimulate and survive exposure to human neutrophils. *J Bacteriol* **195**, 2982-90 (2013).
4. Fox, D.A. et al. Structure of the Neisserial outer membrane protein Opa₆₀: Loop flexibility essential to receptor recognition and bacterial engulfment. *J Am Chem Soc* **136**, 9938-9946 (2014).
5. Dewald, A.H., Hodges, J. C., Columbus, L. Physical determinants of beta-barrel membrane protein folding in lipid vesicles. *Biophys J* **100**, 2131-2140 (2011).
6. Fox, D.A. & Columbus, L. Solution NMR resonance assignment strategies for β -barrel membrane proteins. *Protein Science : A Publication of the Protein Society* **22**, 1133-1140 (2013).
7. Criss, A.K. & Seifert, H.S. *Neisseria gonorrhoeae* suppresses the oxidative burst of human polymorphonuclear leukocytes. *Cell microbiol* **10**, 2257-2270 (2008).

8. Veiksina, S., Kopanchuk, S. & Rinke, A. Fluorescence anisotropy assay for pharmacological characterization of ligand binding dynamics to melanocortin 4 receptors. *Anal Biochem* **402**, 32-39 (2010).