



University
of Glasgow

Nazari Shirvan, Ali (2012) Molecular analysis of surface proteins of *Clostridium difficile*. PhD thesis

<http://theses.gla.ac.uk/3522/>

Copyright and moral rights for this thesis are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the Author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the Author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Molecular analysis of surface proteins of
Clostridium difficile



College of Medical Veterinary and Life Sciences
School of Life Sciences

March 2012

Ali Nazari Shirvan

Supervisor: Dr R. Aitken

A thesis submitted for the degree of
Doctor of Philosophy

Abstract

Since the Gram positive anaerobe *Clostridium difficile* was first isolated and described, it has emerged as an increasingly important nosocomial pathogen in Europe, North America and elsewhere, and the prime causative agent of antibiotic associated diarrhoea and pseudomembranous colitis in humans.

The two large toxins, A and B, are the main virulence factors, proteins that are expressed in the gastrointestinal tract after colonisation by *C. difficile*. The pathological symptoms mediated by these toxins include disruption of the integrity of the the epithelium, fluid loss, intestinal inflammation and tissue destruction.

Important as the toxins are to *C. difficile* pathogenesis, several other proteins are known to contribute to colonisation and other aspects of the disease process remain poorly understood. Immunological studies using antisera from the patients revealed a number of candidates and amongst these, proteins present, or thought to be present, at the bacterial surface contribute to adhesion, motility and other interactions with the human host.

The aims of this study were to produce a number of surface proteins from *C. difficile* as recombinant products and to isolate antibodies against these targets *via* phage display. The goal was to assess if these antibodies could inhibit the normal function of these targets and to confirm their location in *C. difficile*. Of 11 clostridial proteins, expression and purification of 3 proved impossible (Cwp84, FbpA and Acd) but 8 others (CspA, GroEL, FliC, FliD, a putative sortase, Cwp66, and its amino and carboxy terminal regions) were used for antibody isolation along with recombinant and native forms of SlpA. Phage display yielded a large panel of specific single chain variable fragments (scFv) antibodies that were expressed, purified and characterised.

Reaction between the scFvs and their targets took place in ELISA and Western blotting suggesting the recognition of linear rather than conformational epitopes. The binding of scFvs to SlpA and its components showed strain specificity

with good recognition of protein from *C. difficile* 630 but no reaction towards SlpA from R20291, and 027 ribotype. Binding of scFvs of a range of specificities to extracts from *C. difficile* M120 indicated that a component of the S layer from this strain might possess immunoglobulin binding activities in the manner of a superantigen.

The scFvs against flagellar proteins FliC and FliD were able to inhibit bacterial motility and therefore there would be potential in testing whether other scFvs generated in this study were able to inhibit the biological activity of their targets.

Some scFvs were tested in immunofluorescence microscopy. The positive results from these experiments showed that the reagents and the strategy pursued could be used to establish surface exposure of the targets and other components of the bacterial surface. Given the high specificity of the reagents, and in the case of Cwp66, the ability to isolate scFvs against defined regions of the protein, the strategy has the capacity to define the orientation of proteins in the bacterial surface. In contrast, the use of scFvs to locate their targets in electron microscopy using immunogold reagents was unsuccessful. As this approach has been successful in other studies, it deserves further investment of effort.

Overall, expression of proteins from *C. difficile* in an *E.coli* host was generally successful and phage display provided a rapid, highly efficient method for the isolation of specific immunological reagents. These have the potential to explore the location, orientation and activity of proteins from the pathogen.

Table of contents

1	Introduction.....	15
1.1	Overview	15
1.2	Epidemiology.....	17
1.2.1	Europe	17
1.2.2	America	19
1.3	Clinical characteristics	20
1.4	Diagnosis of <i>C. difficile</i>	21
1.5	Risk factors	23
1.6	Prevention and treatment	24
1.7	Prevention and treatment	26
1.8	Life cycle of <i>C.difficile</i>	28
1.9	Pathogenesis	29
1.10	Other virulence factors	32
1.10.1	Surface layer proteins.....	33
1.10.2	Cwp66.....	40
1.10.3	Cwp84.....	41
1.10.4	cwpV
1.10.5	Flagellum	44
1.10.6	Fbp68.....	46
1.10.7	GroEL (Hsp60).....	47
1.10.8	Peptidoglycan hydrolase.....	48
1.10.9	Other surface proteins.....	48
1.11	Surface proteins and the aims of this project	49
1.12	Phage display	50
1.12.1	Principles of phage display	50
1.12.2	Type of libraries and application of phage display	55
1.12.3	Antibody phage display	58
2	Material and methods	70
2.1	Culture of <i>Clostridium difficile</i>	70
2.2	Selection of target proteins.....	71
2.3	Genomic DNA extraction.....	71
2.4	Primer design and PCR	72
2.4.1	Primer design	72
2.4.2	PCR reactions.....	74
2.4.3	Agarose gel electrophoresis	75
2.5	Cloning and expression	75
2.5.1	Overview:	75
2.5.2	Cloning protocol	79
2.5.3	Transformation.....	81
2.5.4	Colony PCR and DNA sequencing.....	81
2.6	Preparation of S-layer extracts using low pH glycine buffer	83
2.7	Flagellar preparation.....	84
2.8	Protein expression and purification	84
2.8.1	Transformation of plasmids into expression strains	84
2.8.2	Protein expression.....	85
2.8.3	Protein purification	86
2.9	Protein electrophoresis.....	87
2.9.1	SDS-PAGE analysis	87
2.9.2	Native PAGE analysis	89
2.10	Western blot analysis.....	89
2.10.1	Western blot.....	90
2.11	Tomlinson library	91
2.11.1	KM13 helper phage	92
2.11.2	Growing the scFv libraries	93
2.12	Selection of recombinant antibodies by phage display	94
2.12.1	Selection on immunotubes.....	94
2.12.2	Screening phage by ELISA.....	96

2.13	Expression and purification of soluble scFv antibodies	98
2.13.1	Expression of solubilised scFv.....	98
2.13.2	Detection of soluble antibodies.....	99
2.13.3	Purification of soluble antibodies.....	99
2.14	Immunochemical characterisation of recombinant antibodies.....	101
2.15	Immunostaining and electron microscopy of <i>C. difficile</i> cells using scFvs	101
2.15.1	Indirect immunofluorescence	102
2.15.2	Electron microscopy	103
2.16	Functional characterisation of anti-flagellar scFv antibodies.....	104
2.16.1	Counting of <i>C.difficile</i> from overnight culture.....	104
2.16.2	Inhibition motility test.....	104
3	Results	106
3.1	PCR product of targets	106
3.2	Cloning and expression of targets by LIC cloning system.....	115
3.2.1	Cloning of the targets.....	115
3.2.2	DNA sequencing analysis.....	119
3.2.3	Expression of the target proteins.....	134
3.2.4	Purification of recombinant proteins.....	139
3.2.5	Mass spectrometry of recombinant proteins.....	153
3.3	Tomlinson Phage display.....	158
3.3.1	Helper phage preparation and titration	158
3.3.2	Preparation and titration phage from libraries (I and J):.....	159
3.3.3	Antibody selection from phage libraries I and J	161
3.3.4	Antibody screening.....	167
3.3.5	Characterisation of selected clones	185
3.3.6	Antibody purification.....	199
3.3.7	Western blot analysis of selected scFvs against targets	204
3.3.8	Analysis of the effect of antibodies on bacterial motility.....	219
3.3.9	Indirect immunofluorescence and electron microscopy.....	222
4	Discussion	228
4.1	Surface proteins in gram positive bacteria	228
4.2	Surface proteins in <i>C.difficile</i>	230
4.3	Recombinant antibodies/ antibody response to different surface protein	236
4.4	Conclusion.....	252
5	Appendices	256
5.1	Brazier's CCEY agar.....	256
5.2	BHI liquid medium.....	256
5.3	Agarose solution in TEA buffer.....	256
5.4	2xYT medium.....	256
5.5	Antibiotic solutions	257
5.6	Western blot solutions.....	257
5.6.1	Electroblotting buffer	257
5.6.2	Ponceau S solution	257
5.6.3	Blocking buffer	257

List of Tables

Table 2. 1. List of sense and antisense primers.....	73
Table 2. 2. PCR reactions.....	74
Table 2. 3. Purified PCR products and their concentrations	80
Table 2. 4. Components of T4 DNA polymerase reaction	80
Table 2. 5. Components of PCR reaction to run the samples for colony PCR	82
Table 2. 6. Solutions of SDS-PAGE separating and stacking gels.....	88
Table 2. 7. Solutions of native PAGE Separating and stacking gels.....	89
Table 3. 1. Sequences of sense and antisense primers for amplification of <i>C. difficile</i> targets by PCR for ligation independent cloning.	110
Table 3. 2. Predicted size and molecular mass weight of the targets.....	111
Table 3. 3. DNA concentrations of extracted samples for use in cloning.....	114
Table 3. 4. Features checked during analysis of cloned target sequences.....	121
Table 3. 5 Titration of KM13 helper phage.....	160
Table 3. 6. Titration of scFv display I and J.....	160
Table 3. 7. Recovery of phage during the selection against CspA.....	162
Table 3. 8. Recovery of phage during the selection against GroEL.....	162
Table 3. 9. Recovery of phage during the selection against FliD.....	163
Table 3. 10. Recovery of phage during the selection against LMW.....	163
Table 3. 11. Recovery of phage during the selection against N-terminus of Cwp66.....	164
Table 3. 12. Recovery of phage during the selection against FliC.....	164
Table 3. 13. Recovery of phage during the selection against C-terminus of Cwp66.....	165
Table 3. 14. Recovery of phage during the selection against Sortase B.....	165
Table 3. 15. Recovery of phage during the selection against Cwp66.....	166
Table 3. 16 Recovery of phage during the selection against slpA.....	166
Table 3. 17 Forward (LMB3) and reverse(gIII) primers used in sequencing of final selected clones.	189
Table 3. 18 Basic sequence of heavy and light chains of scFv which were diversified in the creation of the Tomlinson libraries.....	189
Table 3. 19 Diversified amino acids in CDR2 and CDR3 in heavy chains of sequenced clones.....	192
Table 3. 20 Diversified amino acids in CDR2 and CDR3 in light chains of sequenced clones.	194
Table 3. 21 Comparing of the effect of different scFvs on mobility of <i>C.difficile 630</i>	220

List of Figures

figure1. 1. <i>C. difficile</i> colonies on a blood agar plate.	15
figure1. 3 Diagnostic tests for <i>Clostridium difficile</i>	23
figure1. 4 The typical <i>Clostridium difficile</i> life cycle.	29
figure1. 5 Action of toxins from <i>C. difficile</i> in the intestine.....	32
figure1. 6 Post-translational processing of SlpA.	43
figure1. 7 Formats of engineered antibodies.....	62
Figure 2. 1. LIC strategy showing the structure of the vector, generation and processing of a PCR product, and annealing of vector and insert.	76
Figure 2. 2. Features of the pET-32 EK/LIC vector.	78
Figure 3.1. PCR products obtained with sense and antisense primers using Taq DNA polymerase	108
Figure 3.2. PCR products of targets with Phusion DNA polymerase (New England BioLabs).	109
Figure 3.3. Comparison of PCR reactions using Phusion DNA polymerase with buffers HF and GC buffers, and Taq polymerase.	113
Figure 3.4. Agarose gel electrophoresis of extracted PCR products..	114
Figure 3. 5. Colony PCR of NovaBlue transformants	118
Figure 3.6. Features of the pET-32 EK/LIC expression plasmid that flank the cloning site.	121
Figure 3.7. Alignment shows the intended sequence (“designed”), that recovered from cloning (“r-cwp84”) and the <i>cwp84</i> sequence from <i>Clostridium difficile</i> 630 (“630cwp84”).....	122
Figure 3.8. Alignment of the predicted protein sequence of r-cspA with cspA630 (A) and BLAST output from a search of the protein database (B).....	123
Figure 3. 9. Alignment of the predicted protein sequence of r-fbpA with fbpA630 (A) and BLAST output from a search of the protein database (B).....	124
Figure 3. 10 Alignment of the predicted protein sequence of r-groEL with groEL630 (A) and BLAST output from a search of the protein database (B).	125
Figure 3. 11. Alignment of the predicted protein sequence of r-fliD with fliD630 (A) and BLAST output from a search of the protein database (B).....	126
Figure 3. 12. Alignment of the predicted protein sequence of r-acd with acd630 (A) and BLAST output from a search of the protein database (B).....	127
Figure 3. 13. Alignment of the predicted protein sequence of r-fliC with fliC630 (A) and BLAST output from a search of the protein database (B).....	128
Figure 3. 14. Alignment of the predicted protein sequence of r-5cwp66 and r-3cwp66 with cwp66 630 (A) and BLAST output from a search of the protein database (B and C).	130
Figure 3. 15. Alignment of the predicted protein sequence of r-cwp66 with cwp66 630 (A) and BLAST output from a search of the protein database (B).	132
Figure 3. 16. Alignment of the predicted protein sequence of r-sortaseB with sortaseB630 (A) and BLAST output from a search of the protein database (B).	133
Figure 3.17. Dot blot analysis of the expression of target proteins in <i>E. coli</i> BL21(DE3) in the presence or absence of IPTG.....	135
Figure 3. 18. Western blot analysis of extracts of <i>E. coli</i> BL21(DE3) cells expressing six target proteins	136
Figure 3. 19. Western blot analysis of extracts of <i>E. coli</i> BL21(DE3) cells expressing five target proteins	136
Figure 3. 20. Western blot analysis of extracts of <i>E. coli</i> BL21(DE3) pLysS expressing clostridial target proteins	138
Figure 3.21. Western blot analysis of extracts of <i>E. coli</i> C43(DE3) expressing clostridial target proteins	138
Figure 3. 22. SDS-PAGE analysis of the purification of recombinant CspA	145
Figure 3. 23. SDS-PAGE analysis of the purification of recombinant CspA	146
Figure 3. 24. SDS-PAGE analysis of the purification of recombinant N-terminus Cwp66	147
Figure 3. 25. SDS-PAGE analysis of the purification of recombinant GroEL.....	147
Figure 3. 26. SDS-PAGE analysis of the purification of recombinant FliD	148
Figure 3. 27. SDS-PAGE analysis of purification of recombinant FliD.....	149
Figure 3. 28. SDS-PAGE analysis of purification of recombinant FliC.....	149
Figure 3. 29. SDS-PAGE analysis of purification of recombinant FliC.....	150
Figure 3. 30. SDS-PAGE analysis of purification of recombinant C-terminus Cwp66	151
Figure 3. 31. SDS-PAGE analysis of purification of recombinant sortase B	151
Figure 3. 32. SDS-PAGE analysis of purification of recombinant Cwp66.....	152
Figure 3. 33. Result of Mascot search using data from mass spectrometry analysis of recombinant CspA ..	154

Figure 3. 34. Result of Mascot search using data from mass spectrometry analysis of recombinant N-terminus of Cwp66	154
Figure 3. 35. Result of Mascot search using data from mass spectrometry analysis of recombinant GroEL	155
Figure 3. 36. Result of Mascot search using data from mass spectrometry analysis of recombinant FliD	155
Figure 3. 37. Result of Mascot search using data from mass spectrometry analysis of recombinant FliC	156
Figure 3. 38. Result of Mascot search using data from mass spectrometry analysis of recombinant C-terminus of Cwp66	156
Figure 3. 39. Result of Mascot search using data from mass spectrometry analysis of recombinant Cwp66	157
Figure 3. 40. Result of Mascot search using data from mass spectrometry analysis of recombinant Sortase B	157
Figure 3. 41. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant CspA	169
Figure 3. 42. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant GroEL	169
Figure 3. 43. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant FliD	170
Figure 3. 44. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant LMW	170
Figure 3. 45. Monoclonal phage ELISA of phage from Tomlinson libraries I+J against recombinant N-terminus (left) and C-terminus (right) of Cwp66	171
Figure 3. 46. Monoclonal phage ELISA of phage from Tomlinson libraries I+ J against recombinant FliC (left) and Sortase B (right)	171
Figure 3. 47. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant Cwp66 (left) and extracted SlpA (right)	172
Figure 3. 48. Dot blot analysis of expression of soluble scFv against recombinant CspA	177
Figure 3. 49. Dot blot analysis of expression of soluble scFv against recombinant GroEL	177
Figure 3. 50. Dot blot analysis of expression of soluble scFv against recombinant FliC	178
Figure 3. 51. Dot blot analysis of expression of soluble scFv against recombinant LMW	178
Figure 3. 52. Dot blot analysis of expression of soluble scFv against recombinant N-terminus (a) and C-terminus (b) of Cwp66	179
Figure 3. 53. Dot blot analysis of expression of soluble scFv against recombinant FliC and Sortase B	179
Figure 3. 54. Dot blot analysis of expression of soluble scFv against recombinant Cwp66 and extracted slpA	180
Figure 3. 55. Monoclonal soluble scFv ELISA against recombinant CspA	181
Figure 3. 56. Monoclonal soluble scFv ELISA against recombinant GroEL	181
Figure 3. 57. Monoclonal soluble scFv ELISA against recombinant FliC and Sortase B	182
Figure 3. 58. Monoclonal soluble scFv ELISA against recombinant LMW	182
Figure 3. 59. Monoclonal soluble scFv ELISA against recombinant N-terminus cwp66	183
Figure 3. 60. Monoclonal soluble scFv ELISA against recombinant FliD	183
Figure 3. 61. Monoclonal soluble scFv ELISA against recombinant C-terminus cwp66 and cwp66	184
Figure 3. 62. Monoclonal soluble scFv ELISA against extracted slpA	184
Figure 3. 63. Complete heavy chain nucleic acid sequence with framework regions and CDRs highlighted	190
Figure 3. 64. Light chain nucleotide sequence with framework regions and CDRs highlighted	191
Figure 3. 65. Vector map of phagemid pIT2	196
Figure 3. 66. Sequence region of restriction sites on phagemid pIT2	196
Figure 3. 67. Restriction analysis testing for the presence of full-length scFv inserts in clones from the Tomlinson library	197
Figure 3. 68. Restriction analysis of selected clones from Tomlinson library	198
Figure 3. 71. SDS-PAGE analysis of purified scFvs against recombinant FliC and FliD	203
Figure 3. 72. SDS-PAGE analysis of pooled purified fractions of scFvs against recombinant LMW and extracted slpA	203
Figure 3. 73. Western blot analysis of CspA probed with scFvs	209
Figure 3. 74. Western blot analysis of N terminal of cwp66 probed with scFvs	209
Figure 3. 75. Western blot analysis of groEL probed with scFvs	210
Figure 3. 76. Western blot analysis of fliD probed with scFvs	210
Figure 3. 77. Western blot analysis of fliC probed with scFvs	211
Figure 3. 78. Western blot analysis of N-terminal of cwp66 probed with scFvs	211
Figure 3. 79. Western blot analysis of sortase B probed with scFvs	212
Figure 3. 80. Western blot analysis of cwp66 probed with scFvs	212
Figure 3. 81. Western blot analysis of SlpA extracted from <i>C. difficile</i> 630 probed with scFvs	213

Figure 3. 82 Western blot analysis of SlpA extracted from <i>C. difficile</i> 630, R20291 and M120 probed with scFvs	214
Figure 3. 83. Western blot analysis of SlpA extracted from <i>C. difficile</i> 630, R20291 and M120 probed with scFvs	215
Figure 3. 84 Western blot analysis of flagellum preparation from <i>C. difficile</i> 630 and R20291 probed with scFvs.	216
Figure 3. 85 Western blot analysis of fliC and LMW extracted from <i>C. difficile</i> 630 probed with scFvs.	216
Figure 3. 86 Western blot analysis of cspA from <i>C.difficile</i> , 630, M120 and R20291 probed with scFvs.	217
Figure 3. 87 Western blot analysis of cspA from <i>C.difficile</i> , 630, M120 and R20291 probed with scFvs.	217
Figure 3. 88 Western blot analysis of cspA from <i>C.difficile</i> , 630, M120 and R20291 probed with scFvs.	218
Figure 3. 89 Comparing of different dilution of recombinant antibodies on mobility of <i>C.difficile</i>	221
Figure 3. 90. Detection of flagellum of two strains (630 and R20291) in indirect immunofluorescence test.	224
Figure 3. 91. Detection of surface layer protein on the surface of 630 and M120 strains in indirect immunofluorescence.	225
Figure 3. 92. Immunogold labelling of <i>C.difficile</i> 630 with recombinant antibodies against flagellum.	226
Figure 3. 93. Immunogold labelling of <i>C.difficile</i> 630 with recombinant antibodies against surface layer protein.....	227

Author's Declarations

This thesis is the original work of the author except where otherwise stated.

Ali Nazari Shirvan

Dedication

I dedicate this thesis:

To my beloved parents and family for all of their efforts through my life.

To all people who are involving in human progress in every field, particularly biological sciences, for being and understanding each other in a peaceful world.

Acknowledgements

I would express my best gratitude and most sincere thanks to my supervisor Dr. Robert Aitken for his continued encouragement, supervision and profound enthusiasm, stimulating discussion, patience and reading of, and commenting on this thesis.

I am also grateful to Dr. Gill Douce and her group for their guidance, helpful advice and collaboration during my work.

I wish to thank Mousa Komai, Saber Almadani and all GBRC Staff for their collaboration and advice through this project.

I must sincerely thank my parents, my family and all of my lovely brothers and also friends in Glasgow whose support and encouragement have help me throughout my studies.

Finally, I would like to say many thanks to the Ministry of Health, Iran and Razi Vaccine Institute for my financial support and allowing me to accomplish this project.

Definitions/Abbreviations

BSA: Bovine serum albumin

CH: constant region of heavy chain

CL: constant region of light chain

CDR: complementary determining region

DNA: deoxyribonucleic acid

EDTA: ethylenediaminetetraacetic acid

FR: framework region

GFGE: pulsed field gel electrophoresis

H chain: heavy chain

HRP: horse radish peroxidase

Ig: immunoglobulin

IPTG: Isopropyl- β -D-thiogalactopyranoside

L chain: light chain

LMW: low molecular weight component of SIpA

MAB(s): monoclonal antibody (ies)

MPBS: milk in phosphate buffered saline

PAGE: polyacrylamide gel electrophoresis

PBS: phosphate buffered saline

PCR: polymerase chain reaction

SDS: sodium dodecyl sulphate

SlpA: surface layer protein

ScFv: single chain antibody fragment

V: variable region

VH: variable region of heavy chain

VL: variable region of light chain

1 Introduction

1.1 Overview

Clostridium difficile is a bacteria of the genus *Clostridium*, first described in 1935 and named *Bacillus difficilis* (Hall I, 1935). Since then and based on its emerging importance in the field of human health, much research has been devoted to understanding the pathogenesis of this organism. Despite this effort, the incidence of *C.difficile* infection continues to rise across Europe and America with new outbreaks occurring frequently(Freeman et al.). *Clostridium difficile* is an anaerobic gram positive, spore forming bacteria, 3-5 μm in length and motile like other members of the genus. Typically, colonies are irregular in shape and show optimum growth on blood agar at 37° C after 24-48 h to form colonies 3-5 mm in diameter, grey and opaque in appearance and without obvious haemolytic activity (Figure 1.1).



figure1. 1. *C. difficile* colonies on a blood agar plate.

The bacteria have two forms, vegetative cells and spores, related to the environment around them. Conditions that create stress – the presence of oxygen poor nutrition, aggressive chemicals – trigger the vegetative cells to undergo sporulation so that the organisms can resist hostile situations. *C. difficile* has a wide spread distribution in humans and animals but there is no evidence so far of direct transmission from animals to humans. *C. difficile* is also present in the environment; one study of 2580 samples in the Cardiff area showed these bacteria to be present in rivers (87.5%), beaches (44%), lake water (46.7%), soil (21%), hospitals (20%), and swimming pools (50%). Samples from animal faeces were also positive: assorted farm animals (1%), dogs (10%) and cats (2%) (al Saif and Brazier, 1996).

In a separate study of animal carriage, *C. difficile* was found in the faeces of boiler chickens at market, revealing 29% of birds were positive and of these isolates, 89.7% were toxigenic, although most were also susceptible to antibiotics such as metronidazole and vancomycin (Simango and Mwakurudza, 2008). In another study in Canada, testing of calves revealed that 40% of these animals were positive with 8 different ribotypes recovered, 7 of which are capable of causing infections in humans (Rodriguez-Palacios et al., 2006). These findings are not unusual – similar results have emerged from studies in the United Kingdom (Keel et al., 2007) and the Netherlands (Debast et al., 2009a, Debast et al., 2009b) – and while early isolation of *C. difficile* from different animal sources suggested no relationship with human infection (Oneill et al., 1993), more recently, investigators have questioned this conclusion (Keel et al., 2007, Arroyo et al., 2005). While direct transmission from animals to humans has yet to be demonstrated, spores of *C. difficile* have been isolated from various meat products (Songer et al., 2009, Weese et al., 2009) and vegetables (Bakri et al., 2009).

1.2 Epidemiology

Among the human population, *C. difficile* carriage and infection has been observed at different rates and with different strains since the medical importance of this pathogen first became clear. One of the earliest studies reported the isolation of bacteria from the urinary tract of both males and females (Hafiz et al., 1975). Following shortly after, Larson and colleagues isolated the bacteria from patients with pseudomembranous colitis (PMC) and showed the ability of the organisms to produce toxin *in vitro* (Larson et al., 1978). Early indications of a link with antibiotic usage emerged from a study of patients in postnatal wards which showed a wide range of carriage (2% to 52%) and also a reasonable relationship between antibiotics and the development of PMC (Larson et al., 1982).

Asymptomatic carriage appears to be common. In a study of neonates, a carriage rate of 64% in children aged 1-8 months was revealed (Holst et al., 1981), while this rate was 15.4% in healthy young adults and 7% in healthy elderly adults (Nakamura et al., 1981). Other reports support the finding that neonates typically carry *C. difficile* with high frequency and that very many isolates (71-100%) are toxigenic. To explain these findings, it has been proposed that the toxin receptor is masked in some way in healthy newborn babies (Donta and Myers, 1982, Bolton et al., 1984). Given the increasing numbers of reports of *C. difficile* infection over the years, different countries have devised strategies for managing this infection and following its epidemiology using a range of typing methods.

1.2.1 Europe

Although European countries have adopted different methods for the detection of *C. difficile*, assays for toxin detection and culture based methods remained in common use when Barbut and colleagues carried out an extensive survey 9 years ago (Barbut et al., 2003). Although other methods have picked up in popularity in the intervening period, their benefits over time-consuming reference

methods are not always clear-cut and problems can arise with sensitivity and specificity (Crobach et al., 2009) . Using these and other tools, the mean incidence of hospital-acquired *C. difficile* infection (CDI) was estimated to be 1.1 for 1000 patient admissions. However, it is clear that the pattern of infection is also changing. This is illustrated by an outbreak of CDI in the southeast of England in 2003 – 2004 in which 59% of isolates were typed as PCR ribotype 027, a strain that was unusual for the UK at that time and a strain with similarity to outbreak strains first identified in Canada and USA.

Ribotype 027 strains are hypervirulent and pose a particular threat to elderly patients (Smith, 2005). This is further illustrated by a recent Finnish study that determined that the rate of deaths between 1994 and 2004 that were attributable to CDI increased two-fold for patients over 64 years of age (Lyytikäinen et al., 2009). Strain variation also emerged from a prospective study of 411 suspected isolates that was carried out across 38 hospitals in 14 European countries. In total, 86.1% of all isolates were toxigenic, 322 were positive for the large glucosyltransferase toxins A and B and the majority of them were toxinotype 0. Other toxinotypes such as V, VIII and III were also reported. Toxinotyping is tool for assessing genetic variation at a 19 kb pathogenicity locus (PaLoc) that carries *tcdA* and *tcdB*, the genes for toxin A and toxin B, along with regulatory factors. MIC (minimum inhibitory concentration) determination for 6 antibiotics showed a full susceptibility to metronidazole for all toxigenic isolates. Ribotyping – another molecular tool for strain characterisation that derives data from variation in the 16S – 23S rRNA intergenic region – revealed 6.2% of isolates belonged to the 027 epidemic strain. This strain is characterised by an 18 base pair deletion in the *tcdC* gene that confers hypervirulent properties, and the ability to produce a binary toxin that possesses ADP-ribosyltransferase activity. Mean incidence of *C. difficile* infection was 2.45 per 10000 patients, although this was highly variable between hospitals (Barbut et al., 2007).

The appearance of the 027 ribotype is cause for concern. In a hospital in the Netherlands, an 027 outbreak increased the rate of CDI from 4 cases per 10000 patient to 83 cases in 10000 (Kuijper et al., 2006). Similar patterns occurred in Belgium and Austria. *C. difficile* ribotype 027, toxinotype 111, emerged in seven hospitals in Belgium between 2005-2006 and the incidence of infection increased 3-fold (Kuijper et al., 2006, Indra et al., 2006). Appearance of ribotype 027 has also been reported in Finland, Switzerland, Norway, Hungary, Poland, and Denmark (Pituch et al., 2008, Lytikainen et al., 2007, Fenner et al., 2008, Ingebretsen et al., 2008, Terhes et al., 2009, Bacci et al., 2009, Indra et al., 2009).

1.2.2 North America

There are many reports of the spread of CDI in America and Canada. In America incidence of *C. difficile* colitis showed a significant increase from a base of 0.68% during 1989 and following years to 1.2% in 2000 in hospitalised patients. This proportion of patients with *C. difficile* colitis that went on to develop life threatening symptoms doubled over the same period from 1.6% (1989 through the the latter years of the century) to 3.2% in 2000. Mortality amongst patients that required colectomy also rose sharply (57% in 1989 to 69% in 2000) (Dallal et al., 2002). The results from a large scale study for the period 1991-2003 in Quebec, Canada, revealed 5-fold growth in CDI infection rates from 35.6 per 100000 in 1991 to 156.3 per 100000 in 2003 (Pepin et al., 2004). Data was collected by cytotoxicity assays of patient samples and reports of the causes of death.

Another Canadian report from 12 hospitals documented *C. difficile* infections by the detection of toxins, culture of *C. difficile*, pulse-field gel electrophoresis (PFGE) for strain characterisation, other genetic analysis (binary toxin gene and *tcdC*) and the link with prescription of antibiotics. The investigation found that incidence of infection increased in patients after the age of 50 and that mortality rose after age 60. PFGE results revealed that most strains of *C. difficile*

had the same characteristics as the epidemic 027 strain from the United States (Warny et al., 2005a). The mean of incidence of CDAD in the study was 22.5 per 1000 hospital admissions. Resistance to fluoroquinolones was noted, and the existence of genes for the binary toxin genes and a deletion in the *tcdC* gene, leading to hypervirulent properties, illustrated the new challenges emerging in CDI (Loo, 2006). Spread of the B1/NAP1/027 strain has been reported from across Canada and the USA.

Epidemiologic studies have examined *C. difficile* infection in other parts of the world. A shift of strains was detected in Japan over a 5 year period and in 2007, the 027 strain was isolated from a 30 year old woman after treatment with antibiotics (Sawabe et al., 2007, Kato et al., 2007). Other reports from Korea (Shin et al., 2008, Tae et al., 2009), China (Sundram et al., 2009), and Hong Kong (Cheng et al., 2009) noted the appearance of the 027 ribotype. Although toxigenic strains (078 - 097) have been isolated in Kuwait (Rotimi et al., 2003), there are no documented reports of the 027 ribotype from the Middle East or South American countries. More study and the isolation of greater numbers of samples may be required, following procedures that are now applied in European countries for detection and characterisation of CDI.

1.3 Clinical characteristics

Clinical symptoms of CDI are highly variable from mild to severe diarrhoea and they depend on patient factors such as history of antibiotic treatment or previous exposure to the infection. Treatment with antibiotics and the consequent growth *C. difficile* – either pre-existing in the patient or newly acquired from the environment – that has the capacity to produce toxin (Bartlett, 2002) leads to CDI which can range between mild diarrhoea to pseudomembranous colitis, lower abdominal cramping or pain, fever, nausea and dehydration. Faecal leukocytes and

leukocytosis are two common signs. In fulminant cases, patients may develop toxic megacolon (Kelly et al., 1994).

Renal failure, malignancy, immunosuppression, or the administration of clindamycin are some of the more significant factors that can make patients susceptible to developing severe colitis and symptoms like peritonitis, hyperalbuminemia and alteration in peripheral leukocyte numbers (Rubin et al., 1995). Molecular analysis of bacteria from patients with recurrent disease has revealed that half arise from acquisition of a new strain of *C. difficile* (Oneill et al., 1991). Physical examination such as endoscopy can be used to confirm the existence of white or yellow plaques on the intestinal surface, these features being made up of ruptured cells and mucin, and indicative of the typical signs of pseudomembranous colitis (Mylonakis et al., 2001).

1.4 Diagnosis of *C. difficile*

Since the emergence of CDI, different methods have been proposed for diagnosis of this infection. Most are applied to samples of the patient stool but tests vary widely and include assays for the presence of cytotoxin or glutamate dehydrogenase (GDH) from the bacteria, PCR in different formats, analysis of faecal leukocytes, blood tests, imaging and endoscopy methods (Figure 1.2). Of these, the most reliable include *in vitro* tests for cytotoxin production, toxin-specific immunoassays, assay for glutamate dehydrogenase and RT-PCR. Although diagnosing Glutamate dehydrogenase is not very specific but this assay is very sensitive to find out of the presence of *Clostridium difficile* in stool samples (Goldenberg et al.). Each of these tests requires different times and they also differ in their specificity and sensitivity creating advantages and disadvantages. Among these diagnostic tests, those that take the least time to perform yet report with the highest accuracy are preferred.

New features of *C. difficile* infection such as the production of the binary toxin may require a revision of assay methods, but at present, the cell culture

cytotoxicity neutralization assay (CCNA) is considered the best standard test. This detects the cytopathic activity of the glucosyltransferase toxins A and B based upon their impact upon cultured fibroblast cells and can detect as little as 1pg of toxin in a stool sample. This test is cheap to conduct but relatively time-consuming and comparing results with outcomes from immunoassays for toxin A and B suggests that immunoassays may emerge as a good replacement (Aldeen et al., 2000). An analysis of reports on CDI from 1991 – 2008 revealed a mean sensitivity of 75%-80% and specificity of 97%-98% across a range of popular methods (Crobach et al., 2009).

A new screening test based upon immunoassay of glutamate dehydrogenase can generate results within 24 hours and positive samples can be followed up with the CCNA test (Reller et al., 2007). Of the molecular detection tests, the most sensitive is RT-PCR for detection of toxin B. This shows potential as a replacement for immunoassays (Peterson et al., 2007). Although new commercially-available PCR detection kits are faster and more accurate than the CCNA test, they will need to be updated regularly to keep abreast of the changing nature of CDI as outlined earlier (Swindells et al.).

Question to be answered	Detection method	Advantages	Disadvantages
Is <i>C. difficile</i> present?	Culture	Sensitive, but presence does not equate with infection as many <i>C. difficile</i> strains are non-toxicogenic Useful for epidemiological investigation and surveillance	Slow detection (days) Suboptimal sensitivity in inexperienced hands Requires anaerobic culturing capability
	Antigen (glutamate dehydrogenase) detection	High negative predictive value* Rapid detection (hours)	Not specific for <i>C. difficile</i> and therefore requires supplementary testing
Is <i>C. difficile</i> toxin present?	Cytotoxin assay	Sensitive High specificity for infection	Slow (minimum 1–2 days) Requires access to and/or experience of cell culture methods
	Enzyme immunoassay	Familiar methodology that can be used widely Rapid (hours)	Variable sensitivity and specificity resulting in low positive predictive values, especially in populations with low prevalence of <i>C. difficile</i> infection Requires laboratory facilities
	Membrane assays	Does not necessarily require laboratory facilities Rapid (minutes to hours)	Variable sensitivity and specificity resulting in low positive predictive values, especially in populations with low prevalence of <i>C. difficile</i> infection
Does the <i>C. difficile</i> have the capacity to produce toxin?	Cytotoxigenic culture	High sensitivity	Uncertain specificity for infection Slow (days)
	Detection of toxin B gene	High sensitivity Rapid (hours)	Uncertain specificity for infection Requires laboratory and molecular expertise High cost

*There are recent contradictory data regarding assay sensitivity.

figure1. 2. Diagnostic tests for *Clostridium difficile* (taken from Nature Reviews Microbiology 7, 526-536 [July 2009])

1.5 Risk factors

Many studies have shown that there are three main risk factors for the development of CDI: use of antibiotics, age of the patient and hospitalization. Although other factors such as consumption of acid anti-secretory agents, immunosuppressive treatment and polymorphism in the IL8 gene have been reported, antibiotic therapy consistently emerges as the most important predisposing factor for CDI with clindamycin and quinolones showing particularly strong linkage. One of the earliest studies on the development of *C. difficile* infection amongst patients taking clindamycin revealed 21% developed diarrhoea and 10% progressed to pseudomembranous colitis (Tedesco et al., 1974). Many

other reports have shown an association between clindamycin use and *C. difficile* outbreaks and mutations in *erm* gene(s) encoding a 23S ribosomal RNA methylase (Johnson et al., 1999).

Large studies have also shown that exposure to fluoroquinolones is also linked with CDI outbreaks and the reverse, that limiting use of these antibiotics is an effective way to decrease the frequency of outbreaks (Muto et al., 2005). An *in vitro* study of an outbreak in Quebec used pulsed field gel electrophoresis (PFGE) to discriminate between *C. difficile* strains and revealed that all were susceptible to such antibiotics such as vancomycin but resistant to ceftriaxone, clarithromycin, gatifloxacin and moxifloxacin (Bourgault et al., 2006). Changing from use of gatifloxacin to levofloxacin decreased the rate of CDI (Gaynes et al., 2004). This study and others reinforce the idea that control of antibiotic choice and variation of the drugs that are used can be effective in decreasing rates of CDI.

Turning to the importance of anti-secretory drugs, a cohort study conducted in Montreal revealed that proton pump inhibitors and H2 blockers were potential risk factors and that they should be used with care for the patients at threat of CDI (Dial et al., 2004). This was supported by data from the United Kingdom (Dial et al., 2005). Immunosuppression and IL8 polymorphism are other risk factors which have been reported (Sanchez et al., 2005, Jiang et al., 2006).

1.6 Prevention and treatment

There are two main ways to prevent CDI: careful use of antimicrobial agents and blocking transmission. To minimize the effect of antimicrobial agents, patients should be treated with drugs which have minimal effect on the intestinal microflora. Transmission can be blocked by having a managed system for the care of patients that have developed CDI, maintaining isolation and exercising meticulous hygiene. Although antibiotics remain mainstay of treatment for CDI,

other approaches are under development such as immunotherapy, probiotics, vaccination and polymers that can adsorb bacterial toxins and other virulence factors.

Vancomycin and metronidazole are the main drugs used to treat CDI (Gerding, 2005). A study of 27 patients showed that metronidazole taken orally for 7-10 days was effective, but could fail if co-administered with the antibiotic that originally triggered CDI (Modena et al., 2006). Analysis of relapse of CDI revealed vancomycin was better than metronidazole in removing *C. difficile* as assessed by culture or the presence of toxin in the patient stool. When vancomycin was delivered in pulsed doses, recurrence of infection was significantly decreased (McFarland et al., 2002). More recent studies have reported reduced efficacy with metronidazole treatment and in a study of 207 patients treated with metronidazole, *C. difficile* was cleared from 50% but 22% underwent relapse and symptoms of colitis returned. The emergence of resistance to metronidazole is an issue of concern (Musher et al., 2005).

In addition to antibiotics, prebiotics and probiotics are under investigation as treatments for a range of gastrointestinal disorders including *C. difficile*. Use of vancomycin has been explored, with or without *Saccharomyces boulardii* as a probiotic. Decreased relapse rates were reported in patients that received vancomycin and the probiotic, when compared to groups that were treated with the antibiotic alone (Surawicz et al., 2000). Using the antibiotic with oligofructose as a prebiotic, also showed benefits (Lewis et al., 2005). Tolevamer is a novel toxin binding polymer that can be used to treat *C. difficile*. In a hamster model, the polymer reduced the severity and recurrence rate of CDI and in the same study, the enterotoxic effects of toxin A were reduced by administration of tolevamer to rats (Barker et al., 2006).

Oral immunotherapy or passive intravenous immunoglobulin (IVIG) have also been considered as strategies against *C. difficile* infection and there are many

reports showing the effect of passive immunity in treatment of CDI. Immunized female hamsters can transfer protection to their offspring (Kim et al., 1987) and in other work with this animal model, concentrated hyperimmune bovine IgG, prepared from cattle vaccinated with toxin A and B, was able to protect against CDI providing treatment commenced before the onset of diarrhoea (Lyerly et al., 1991). In a study with human patients, oral delivery of whey protein concentrates from cattle immunised with inactivated toxins and whole cell killed *C. difficile* proved effective (van Dissel et al., 2005). These whey protein concentrates are particularly rich in bovine secretory IgA and appear safe to use with CDI patients (Young et al., 2007). Passive intravenous therapy with human immunoglobulin also shows potential in the limited number of reports in the literature (Abougergi and Kwon).

Vaccination with inactivated toxins from *C. difficile* is also another field of immunology relevant to the control of this infection. Reports of the use of parenteral vaccines containing toxoid A and B or separately purified proteins showed a strong response to toxin A and protection against CDI. The results from some studies show no recurrence of CDI after vaccination, strongly supporting the view that this method has potential against CDAD and colitis (Sougioultzis et al., 2004, Aboudola et al., 2003, Giannasca and Warny, 2004b).

1.7 Typing of *Clostridium difficile*

There are different typing methods for *C. difficile* can be divided into two groups, phenotypic and genotypic methods. In phenotypic methods the characteristics of the surface components of bacteria are used in typing schemes. Typing *via* this method can use different technical approaches such as SDS-PAGE or the immunochemical fingerprinting of cell extracts. Combination of these methods showed some value in documenting the differences between isolates in an

outbreak but methods based on the genotype would be useful for cosmopolitan typing of *C. difficile* (van den Berg et al., 2007).

The earliest methods of serotyping used antisera which were prepared against formol treated strains. Serogrouping of isolates through slide agglutination revealed 6 groups which differed in toxin production and association with antibiotic-associated diarrhoea, serogroups A, C and D being identified in cases of antibiotic-associated diarrhoea (Delmee et al., 1985). Another study reported the development of 15 serogroup antisera which were used to identify 6 cytotoxigenic serogroups isolated from mildly to severe cases of CDI (Toma et al., 1988).

While different studies used these methods to interpret the behaviour of new isolates, molecular methods based on bacterial genotyping have been developed. In summary, PCR ribotyping (analysis of the spacer regions between 16S and 23S ribosomal RNA genes), pulsed field gel electrophoresis (PFGE), multilocus variable number tandem repeat analysis (MLVA), restriction endonuclease analysis (REA) and toxinotyping are the most frequently reported. They have all their advantages and disadvantages. Activity since the 1980s has aimed to improve typing methods and to select the best, and from this, molecular methods have emerged because of the stability of the clostridial genome and the reproducibility of the methods (Brazier, 2001).

For quick recognition of strains, REA, arbitrary primer PCR and PCR ribotyping are preferred over MLST or MLVA which are better suited to long term epidemiology studies (Kuijper et al., 2009). The new outbreak strain of *C. difficile* has been variously termed B1/NAP1/027 based on different typing methods: by REA, it belongs to the B1 group; by PFGE, it belongs to North American PFGE type 1 (NAP1); by ribotyping it belongs to ribotype 027. The toxinotyping method places the strain in toxinotype III, a group that in addition to toxin A and B, shows the existence of the binary toxin CDT and deletion of 18 base pair in *tcdC*, the gene for

a negative regulator of toxin A and B. These properties help explain the virulence of the strain.

Although this strain is relatively new and different from the strains first reported in 1984, differences have been noted in its incidence and resistance to fluoroquinolones (McFarland et al., 2007). In a study that analysed *C. difficile* from 124 patients, isolates were characterised by different typing method that examined the deletion in *tcdC*, production of toxins A and B, and the presence of the binary toxin. Results showed B1/NAP1/027 produced 16 times more toxin A, and 23 times more toxin B than was typical (Vohra and Poxton, Vohra and Poxton, Warny et al., 2005b).

1.8 Life cycle of *C. difficile*

The ability of *C. difficile* to form spores poses a particular challenge in control of the pathogen, particularly in the hospital environment where rigorous cleaning methods are required to contain spread (Figure 1.3). Spores can be transmitted by inhalation or ingestion. Gastric pH provides effective protection against the vegetative cells of many pathogens but the ability of spores to germinate can also be effected to some extent (Fordtran, 2006). In the lower part of the gastrointestinal tract, germination is triggered by factors such as heat, pH, bile acids, and glycine (Wheeldon et al., 2008). Proton pump inhibitors increase the rate of germination, enhancing the patient's risk of contracting *C. difficile* infection (Nachnani et al., 2008, Dalton et al., 2009) (Figure 1.3). Germination can be rapid – 78% of spores can germinate in the small intestine one hour after reaching this destination (Wilson et al., 1985).

Following germination of spores in the intestine, *C. difficile* does not appear to pose a significant typo as the normal microflora inhibits its colonization. Under some conditions (broad-spectrum antibiotic therapy, long term hospitalization,

surgery or other circumstances that disrupt the normal flora) the vegetative cells of *C. difficile* can colonize and under these circumstances, the characteristic symptoms of infection can emerge (Figure 1.3). The development of new diagnostic methods and awareness of the potential consequences of antibiotic therapy has done little to curb the rise of *C. difficile*-associated diarrhoea (Elliott et al., 2007).

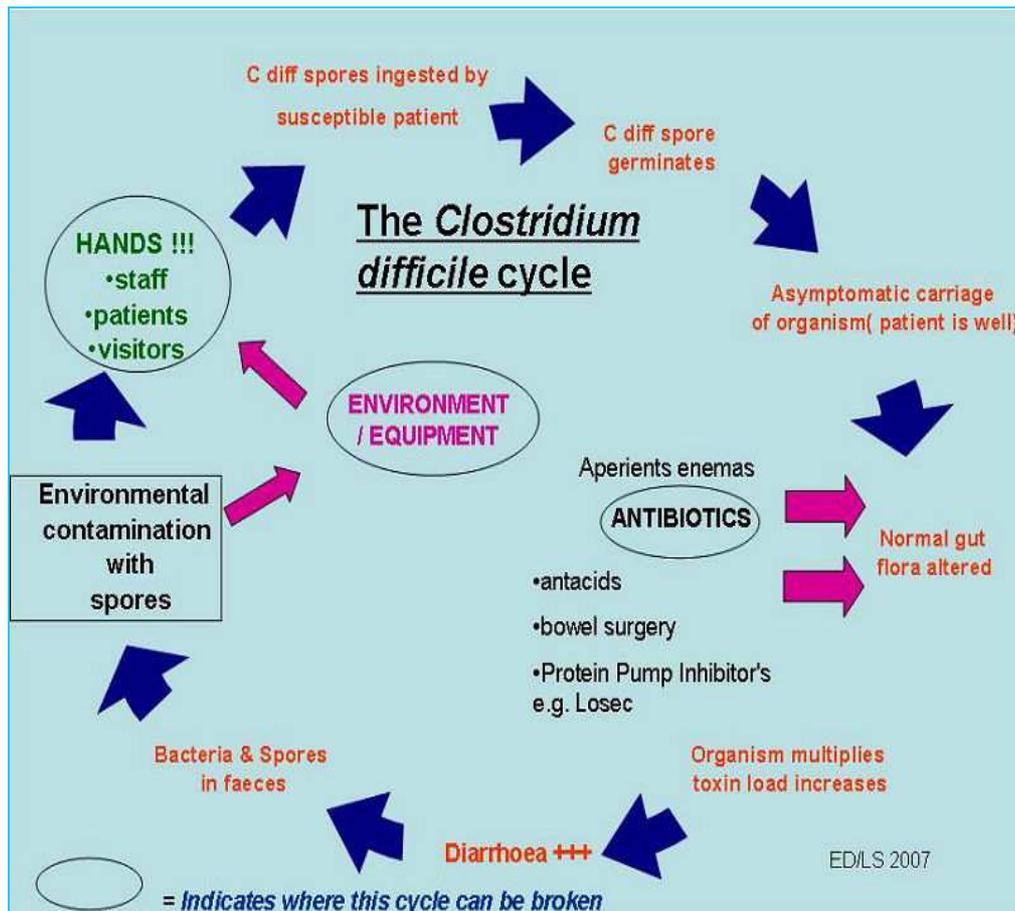


figure1. 3. The typical *Clostridium difficile* life cycle.

Taken from Southend University Hospital web site

1.9 Pathogenesis

Pathogenesis of *C. difficile* infection is a multi-step process. The first step is the creation of the appropriate microenvironment for germination of spores in the gastrointestinal tract. Spores, which can be acquired from the external environment or from existing clostridial cells in the intestine, germinate and new

vegetative cells then establish through active colonisation. Penetration of the mucus layer and attachment of clostridial cells to the epithelium takes place through the activity of surface factors such as surface layer proteins and flagella but current understanding of these processes is incomplete. The production of toxin A and B, the main virulence factors, then ensues.

The pathogenicity locus (PaLoc) on the *C. difficile* genome carries the genes for toxin A and B. This 19.6 kb locus comprises five genes for the regulated production of the toxins (*tcdABCRE*). *TcdA* and *tcdB* genes are the toxin structural genes. There are two regulator genes: *tcdD* is a positive regulator; *tcdC* is a negative regulator. *tcdE* encodes a protein which may play a role in releasing the large toxin molecules from the bacterial cell. The first strain to be sequenced, strain 630, can produce both toxin A and B. Several other strains possess a third toxin known as the binary toxin which is encoded near to the PaLoc (Sebahia et al., 2006)

Both toxins A and B are members of a family of large clostridial toxins. Based on variations in the genes for these toxins and elsewhere in the PaLoc, there are more than 20 toxinotypes. Toxin A (308 kDa) and B (270 kDa) comprise three domains. At the amino-terminus, they have an enzyme domain and the carboxy-terminus is a region responsible for receptor binding, while the central part with hydrophobic amino acids is believed to be involved in translocation of the toxin into the cytosol of host cells (Jank et al., 2007). Sequence and crystallographic analysis of the carboxy-terminus showed there are repeated sequence in this domain and through these repeats toxins, the toxins can bind to cell surface carbohydrates (Ho et al., 2005).

Following receptor binding, the toxins are internalised by endocytosis and after that, two steps are required for the toxins to exert their effects. First is the translocation of the toxins into cytosol. This is thought to be induced by the progressive acidification of the endocytic vesicle which induces translocation of

toxins from the endosome interior into the cytosol (Barth et al., 2001). Once in this compartment, the enzyme activity of the toxins drives glucosylation of Rho GTPases. The resulting inhibition of the activity of Rho proteins causes down regulation or inactivation of a wide range of vital cell activities such as actin cytoskeleton regulation, epithelial barrier function, cell deaths, wound repair and phagocytosis, reflecting the involvement of Rho GTPases in many aspects of cell biology (Jank et al., 2007).

There are many reports showing the multiple effects of toxins A and B can lead to disruption of intestinal function, proceeding to pseudomembranous colitis (Voth and Ballard, 2005) (Figure 1.4). The advent of gene knockout systems for *C. difficile* (Heap et al., 2007) has allowed the relative contributions of toxins A and B to pathogenesis to be assessed in animal models. This has revealed that toxin B is more important in virulence than toxin A (Lyras et al., 2009) and that its damaging effect may be 10 times greater than that of toxin A (Riegler et al., 1995). Production of the toxins is influenced by different environmental signals. Depending on the strain, nutrient levels, temperature and sub-inhibitory levels of antibiotics can all contribute (Freeman et al., 2005) and a modest depletion of biotin (0.05 nM) results in higher production of toxin A (64 fold increase) and toxin B (35 fold increase) (Yamakawa et al., 1996).

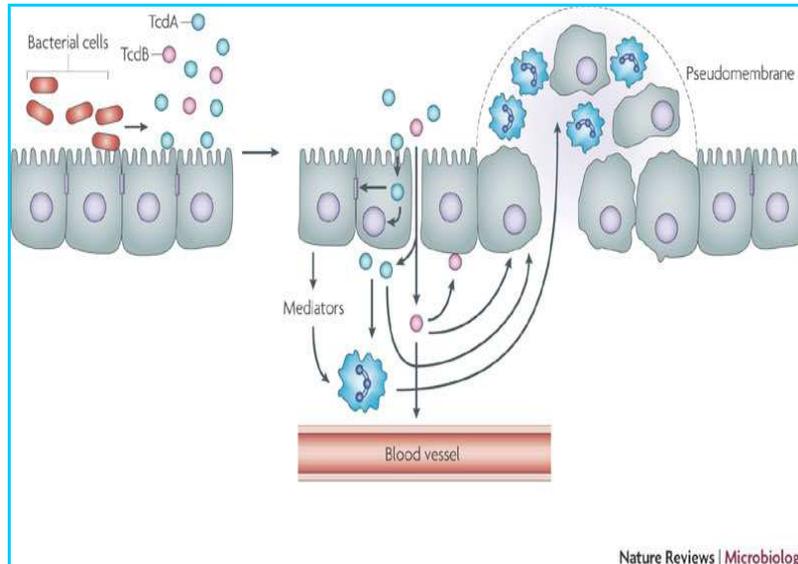


figure1. 4. Action of toxins from *C. difficile* in the intestine.

C. difficile cells are shown free and attached to the intestinal epithelium as red shapes. TcdA and TcdB are shown as coloured circles. Binding of TcdA to apical surfaces causes shape change, disruption of tight junctions and loosening of the epithelial barrier. This in turn allows TcdB to bind basolateral cell surfaces. Inflammatory mediators are induced by toxin action and attract neutrophils leading for formation of the pseudomembrane. Taken from Nature Reviews in Microbiology 7, 526-536 (July 2009).

1.10 Other virulence factors

Adherence to the intestinal wall is likely to be a necessary step in the infection process before toxin production. *C. difficile* has been shown to possess an ability to attach to different cells *in vitro* (eg Caco-2 and Vero cells) and *in vivo*, adherence to the caecal mucus of mice has been reported. In this animal study model, germ free mice were used and reproducible adherence to the caecum was observed, a process that was compared to adherence *ex vivo*. Data was reproducible for most strains, adherence to mucus being better for toxigenic strains than non toxigenic strains (GomezTrevino et al., 1996).

Analysis of bacterial adherence to Caco-2 cell monolayers revealed that with all *C. difficile* strains tested, binding mainly took place at the outer edge of cell clusters. Use of chelating agents increased bacterial adherence and binding was noted at the basolateral surface of cells and directly to protein fibres of the extra cellular matrix. This study suggested two processes for adherence, the first being direct attachment to a receptor on the lateral part of Caco-2 cells and the second

via binding to the extracellular matrix (Cerquetti et al., 2002). One can imagine a scenario *in vivo* in which adherence of bacteria to the apical surface of the human intestinal epithelium, followed by toxin production, would disrupt the integrity of the wall and allow further rounds of adherence to take place at the basolateral surface of cells and to components of the extracellular matrix.

More detailed studies of virulence factors and other proteins present at the surface of *C. difficile*, proteins such as SlpA, Cwp66, Cwp84, FliC, FliD, GroEL, and the fibronectin binding protein, has the potential to help clarify the substantial uncertainties of the process of colonisation and other aspects of the pathogenesis of this pathogen.

1.10.1 Surface layer proteins

Surface layer proteins are found in a number of different bacteria. These proteins are involved in various activities such as the formation of a protective layer, cell adherence to surfaces, the creation of molecular sieves, as a scaffold for attachment of other virulence factors. They have a variety of molecular weights (40-170 kDa) with the capacity to interact with specific glycan chains and ions. Their location and their diverse functions makes a case for exploitation in vaccine development, and their ability to spontaneously self-associate into layers has been used in nanobiotechnology (Debabov, 2004).

Analysis of serum IgG from patients with the symptoms of CDI showed a positive reaction with a protein of around 36 kDa present in EDTA extracts of clostridial cells, showing that surface proteins are the target of immune responses in patients (Pantosti et al., 1989a). The 36 kDa protein was extracted from *C. difficile* C253 and purified. Results showed it could adopt a dimeric structure under native conditions but that it lacked glycan content. Immunofluorescence analysis made clear that this antigen is exposed on the surface of bacterial cells and has

similarity with S layer proteins from other bacteria. It lacked toxic or protease activity (Cerquetti et al., 1992a).

Electron microscopy of the cell wall of *C. difficile* GAI0714 showed a two layer structure with a 20 nm inner layer and a 10 nm outer layer. Removal of the outer layer completely changed the normal regular array present at the bacterial surface. Although the study showed a square array of components, determination of the true structure was difficult because of fixation conditions and autolysis. SDS-PAGE analysis of cell wall preparations showed a range of proteins of 45-47 kDa and 32 kDa were extracted from some strains of *C. difficile* while the others showed proteins of 42 and 38 kDa (Kawata et al., 1984).

Result from reassembly experiments using the main two proteins extracted from bacteria showed the native array pattern could be reconstructed when the proteins were dialysed against CaCl₂. The authors concluded that the outer cell wall layer of this pathogen comprised two proteins that required the presence of Ca²⁺ for formation of an array (Masuda et al., 1989).

Further investigations suggested that both components of the S layer – in *C. difficile* GAI0714, proteins of 32 and 45 kDa – can as dimers. Analysis showed that the two proteins were different in amino acid composition, but both were composed mostly of acidic amino acids, without cysteine and with low amounts of histidine and methionine. On account of their composition, pI values of 3.7-3.9 were determined for the low molecular weight protein and 3.3 for the high molecular component of the S layer (Takeoka et al., 1991).

Another report showed that the major surface proteins are variable in molecular weight when extracted by urea and also that they are strain specific with the exception for 73 kDa antigen present in more modest quantities in the urea extracts (Sharp and Poxton, 1988). In related work (Cerquetti et al., 2000b), SlpA from six clinical isolates was characterised by a range of different methods. Electron microscopy data confirmed the presence of a regular crystalline array on

the surface of isolates. Although a lattice pattern could be recovered *in vitro* using urea extracts from cells, this was not possible when attempted with purified proteins either as single components or mixtures of the two dominant proteins present in urea extracts. These two proteins possessed molecular weights in the range 36-56 kDa and from all strains tested, they appeared to be glycoproteins. Rabbit antisera were raised against the 36 and 47 kDa proteins of strain C253 and used to understand better their antigenic relationship with components from the other six strains. Western blots showed antisera against the 47 kDa protein from C253 could react with high molecular weight SlpA proteins from the other six strains. In contrast, antisera to the 36kDa protein of C253 failed to recognise the low molecular proteins from other strains. Amino terminal sequencing revealed some similarity between low and high molecular parts of SlpA (Cerquetti et al., 2000b).

A further study was done to characterise and compare the *slpA* genes and their proteins between different strains and ribotypes. Low pH extraction of SlpA from ribotypes 1 and 17 and also from strains Y and 630 showed a high molecular weight component of 41-45 kDa and 33-38 kDa for the smaller protein. While native electrophoresis of extracts from ribotypes 1 and 17 produced bands in excess of 140 kDa, other methods confirmed earlier reports of the formation of dimers or tetramers of SlpA components. Sequence analysis revealed that high and low molecular weight proteins are derived by post-translational cleavage events, the first removing the signal peptide from the amino terminus of the protein, the second taking place at an internal site.

Following removal of the signal sequence, the low molecular weight protein of SlpA derives from the amino terminal region and the carboxy terminal region constitutes the high molecular protein. Alignment of protein sequences for SlpA from ribotypes 1, 17 and strain 630 showed a conserved signal peptide and high similarity between mature forms of the high molecular weight proteins. In contrast, low molecular weight proteins were non identical (Calabi et al., 2001).

Analysis of genomic data from strain 630 showed that *C. difficile* possesses a number of open reading frames (ORF) with homology to the high molecular weight form (HMW) of SlpA and also similarity between the HMW and an autolysin from *B. subtilis*. Immunological analysis showed that the HMW is immunologically conserved between *C. difficile* strains in contrast to the diversity of LMW (Calabi et al., 2001). Working with the genome of *C. difficile* 630, the *slpA* gene (2160 bp) was predicted to encode a 73.4 kDa protein, the precursor for HMW and LMW.

Further sequence analysis of PCR products showed the gene to be similar between strains, C253 and 630. It was revealed that SlpA has 78 amino acids at its amino terminus with homology to an SLH domain, a feature detected in many S-layer proteins and thought to be responsible for anchoring to the bacterial surface. In the carboxy terminal part of SlpA, homology was noted with a domain with potential of amidase activity. More analysis determined a 37 kb region of the clostridial genome with 17 ORFs of which 11 ORF carried domains with homology to the cell wall anchoring domain of *cwlB* autolysin *B. subtilis*.

In almost all cases, these conserved domains lie on ORFs with a signal peptide and another variable domain with a range of functions (*eg* protease activity, Cwp84) (Karjalainen et al., 2001). Although deeper analysis of the S-layer proteins showed again a huge variation in the LMW between strains, conservation of a site for cleavage of LMW from HMW was observed consistently. The high rate variation in LMW sequences suggested that this might serve as means to evade the immune response of the host (Calabi and Fairweather, 2002).

Analysing of the variable part of the *slpA* gene by PCR-RFLP and nucleotide sequencing showed this region is similar within a serogroup but few similarities exist between different serogroups. Although this analysis showed that the *slpA* gene can be a target for discrimination of *C. difficile* isolates, sequencing of more strains is required (Karjalainen et al., 2002). In Japan, PCR of the *slpA* gene of strains of the smz ribotype (Kato et al., 2001) has been used along with other

methods (tests for toxin A and B, PCR ribotyping and PCR for *tcdA*) as part of a typing scheme. Analysis of 10 smz strains revealed that three subtypes were detectable (smz-1, smz-2 and smz-3) on the basis of *slpA* sequencing (Kato et al., 2005). This typing method has been applied in different wards of a hospital using the stool specimens from patients with diarrhoea. Two smz-1 and smz-2 subtypes were found amongst major cases (Kato et al., 2009). Analysis of endemic isolates in Japan revealed this typing method can be used in conjunction with testing for toxin A, B and binary toxin and also ribotyping with good reproducibility. Based on *slpA* sequencing, 87 isolates were divided into 14 main types and 18 subtypes, an outcome that was compatible with PCR ribotyping of isolates (Kato et al., Kato et al.).

Despite sequence differences, expression *slpA* and processing of its translation products seems to follow a consistent pattern. Signal sequences comprise the amino terminal 24 residues and cleavage has been shown to occur at the last amino acid of a conserved sequence, SAAPVFA. The second predicted cleavage site was found to have an absolute requirement for a conserved GKR motif close to the cleavage site and often takes place amino terminal to a serine residue, resulting in a TKS or TYX motif at the carboxy terminus of the LMW product. The TAA stop codon was found in *slpA* in most ribotypes except ribotype 078 which was terminated with TAG.

Downstream of the stop codon, sequence typical of Rho-independent transcriptional termination was found. The *slpA* gene was found to lie close to *secA* with the two genes separated by 202 to 268 bp depending on ribotype (Eidhin et al., 2006). Analysis of *slpA* genes from two *C. difficile* strains of ribotype 027 (CD196 and R20291) showed that both have cleavage sites after amino acids 24 and 342, to generate HMW and LMW proteins of 44 and 34 kDa respectively.

Comparison of the amino acid sequences of *slpA* indicated high similarity between ribotype 027 and the epidemic ribotype 001 (Spigaglia et al.).

Although it is very likely that binding of *C. difficile* to the intestinal epithelium is multifactorial, the amino acid composition and high representation on the bacterial surface suggests that SlpA may have a contributory role and it has been shown that both native and recombinant SlpA can bind to HEp-2 cells *in vitro*. Acidic extraction of SlpA from three strains (1, 17 and 630) showed this property and that it was independent of the bacterial strain. Analysis of recombinant HMW and LMW showed binding of HMW to host cells was the more effective and this was confirmed by using fluorochrome-labelled *C. difficile* and specific antisera in a blocking assay with human epithelial cells.

In a further step, the investigators tested for binding of native and recombinant S layer proteins using gastrointestinal tissue sections. They noted that HMW could both bind strongly to the epithelium and the lamina propria of mucosa in all level of gastro tract, while LMW gave a much weaker signals and this was limited to the epithelium. Blot analysis also described the binding of recombinant HMW and acidic extracts from strain 630 to collagen I, thrombospondin and vitronectin which are some of the components of the extracellular matrix (Calabi et al., 2002).

While different studies have described the presence of a glycan chain on S-layer proteins of *C. difficile* (Calabi et al., 2001, Cerquetti et al., 2000a, Mauri et al., 1999), recent research has demonstrated that this protein is not glycosylated. This conclusion was reached from analysis of different S-layer proteins from different strains using mass spectrometry. The authors propose that a lack of reliability in commercial glycan-labelling kits and contamination have generated misleading results (Qazi et al., 2009). Other improvements to understanding of the S-layer protein have come from structural analysis of the interaction between HMW and LMW. Construction of different derivatives of both these proteins has shown the

importance of residues 260-312 of LMW and the first 156 residues of HMW to the ability of these two proteins to interact. Crystallographic analysis of a truncated LMW (residues 1-262; native length 321 residues) has revealed a novel protein with two domains: domain 1 comprises residues 1-88 and 239-249, while domain 2 is formed from residues 82-238. The investigators have proposed an end-to-end complex between HMW and LMW with HMW facing the bacterial cell wall and LMW facing the external environment (Fagan et al., 2009).

S-layer proteins are targeted by the host immune response and potentially, LMW could be more immunogenic (Cerquetti et al., 1992a). A study of the human antibody response to surface layer protein failed to find substantial differences between serum IgM, IgG and IgA levels at times post-exposure among CDI patients, asymptomatic carriers and controls. However, it has been noted that high levels of IgM against surface layer proteins was accompanied with decreasing recurrence of CDAD which may reflect the contribution of this protein to the process of colonisation (Drudy et al., 2004). Linking with this, passive immunisation of hamsters with antiserum raised against purified S-layer proteins showed extended survival time after challenge with *C. difficile* compared to controls. This was the basis for a suggestion that SLPs might form components of a vaccine (O'Brien et al., 2005).

Amongst other activities, SLP can also alter the function of monocytes and dendritic cells (DC) which are important in innate and immune defence. Monocytes treated with purified SLP from *C. difficile* C253 showed increases in the production of the pro-inflammatory cytokines IL-6 and IL-1 β that were comparable to those stimulated by LPS from *E. coli*. The maturation of monocyte derived dendritic cells can be followed by markers such as CD83 and MHC Class II. Treatment with SLP from C253 provoked expression of these markers and the production of the pro-inflammatory cytokine IL-12 p70 and the anti-inflammatory cytokine IL-10, activities that can induce orientation of the immune response to Th1 and Th2 respectively (Ausiello et al., 2006). In further work, the

immunomodulatory properties of SLP from different strains were considered. No differences were noted between hypervirulent epidemic strains and non-epidemic strains in the capacities of SLP to stimulate production of IL-1beta and IL-6 or cell markers. Some differences were noted in the stimulation of IL-10 and IL-12p70 (Manuela Bianco, 2011). It has also been noted that the LMW component of SlpA has some adjuvant properties (Brun et al., 2008).

Based on analysis of the genome of *C. difficile* 630 (Sebahia et al., 2006), it is predicted that 29 genes carry three copies of the cell wall binding Pfam 04122 motif (Fagan et al.). Examples include Cwp66 and Cwp84, as described below. Proteomic analysis of surface extracts has shown the presence of many other proteins (Wright et al., 2005) but the abundance, sequence variation, and multiple activities of the HMW and LMW component of SlpA, along with the bias of patient responses towards these proteins (Wright et al., 2008) lends them a special importance in understanding the pathogenesis of CDI.

1.10.2 *Cwp66*

Studies on the ability of *C. difficile* to bind to different cultured cell lines (*eg* Vero, Hela) discovered that after a heat shock at 60°C, bacterial binding was increased. Analysis identified a surface protein that appeared important in the adhesion process (Karjalainen et al., 1994). Iron limiting media, high solute concentrations (*eg* NaCl₂, 550mM; CaCl₂, 25mM) and acidic shock were also shown to enhance adhesion. Three surface proteins of 70, 50 and 40 kDa were implicated in this adhesion process (Waligora et al., 1999).

These studies lead to the identification of *cwp66*, a gene lying downstream from *slpA*. Two open reading frames were located, *orfA* and *orfB*, that are separated by a 32 bp gap. They appeared to comprise a small operon with a *rho*-independent termination sequence. The investigators noted that *orfA* (651 bp) was

responsible for production of a 24.9kDa protein while *orfB* (1830 bp) encoded Cwp66, the surface-located adhesin. Cwp66 is made up of 610 amino acids, forming a protein of 66 kDa with a pI of 5.3. Primary and secondary prediction showed Cwp66 to be a hydrophilic protein with two domains. Comparison between strains indicated variability in the carboxy terminal part and conservation in amino terminal region. As with SlpA, similarity was noted with the autolysin CwlB of *B. subtilis* at the amino terminus, suggesting this region of Cwp66 has a role in attachment of the protein to the bacterial cell wall.

Examination of heat-shocked bacteria using immunogold and specific antibodies against recombinant amino and carboxy terminal parts of Cwp66 revealed that the carboxyl terminal domain is a surface exposed. The antisera also showed some capacity to inhibit the adherence of *C. difficile* to cultured cell lines (Waligora et al., 2001). PCR-RFLP with different strains showed heterogeneity in *cwp66* in the 3' region (Pechine et al., 2005b).

1.10.3 Cwp84

Located just down stream of *cwp66*, *cwp84* encodes another protein that is located at the bacterial surface. *cwp84* was determined to be monocistronic and to be expressed early in exponential phase of growth. A cell-anchoring domain appeared to be present at the carboxy terminus while the amino terminal part of the protein possessed the features of a cysteine protease (Savariau-Lacomme et al., 2003). Deletion mutagenesis confirmed the location of the enzymic domain of Cwp84 but purification of the recombinant protein was unsuccessful possibly because of putative autocatalytic activity which has been reported for other cysteine proteases. The sequence of the functional domain was found to be highly conserved with little or no variation among different serotypes of *C. difficile*, suggesting an important role for this protease, possibly in the assembly of proteins at the bacterial surface (Savariau-Lacomme et al., 2003).

Further study characterised the biochemistry of Cwp84 and its proteolytic effect on the host extracellular matrix (ECM). SDS-PAGE analysis of purified recombinant Cwp84 using a fused histidine tag protein confirmed autocatalytic activity under reducing conditions leading through step-wise changes, to a mature protein of 61 kDa. ECM proteins such as fibronectin, laminin and fibronectin can be cleaved by Cwp84 under reducing conditions, and protease activity was inhibited by E64 or anti-Cwp84 antibodies (Janoir et al., 2007). The protease could be detected in surface extracts and was expressed by different clinical strains. Antibody responses against Cwp84 could be detected in patient sera showing its expression and immunogenicity *in vivo*. It has been suggested that the protein might be a good vaccine candidate (Pechine et al., 2005b).

More recent reports have shown a role for Cwp84 in the maturation of surface layer proteins in *C. difficile*. Bacteria carrying deletions of *cwp84* had different colony morphology and were slow growing in liquid culture with a tendency to aggregation. In *cwp84* mutants, SlpA processing was incomplete but results from an animal model of *C. difficile* infection showed no substantial differences in pathogenesis compared with wild type (Kirby et al., 2009).

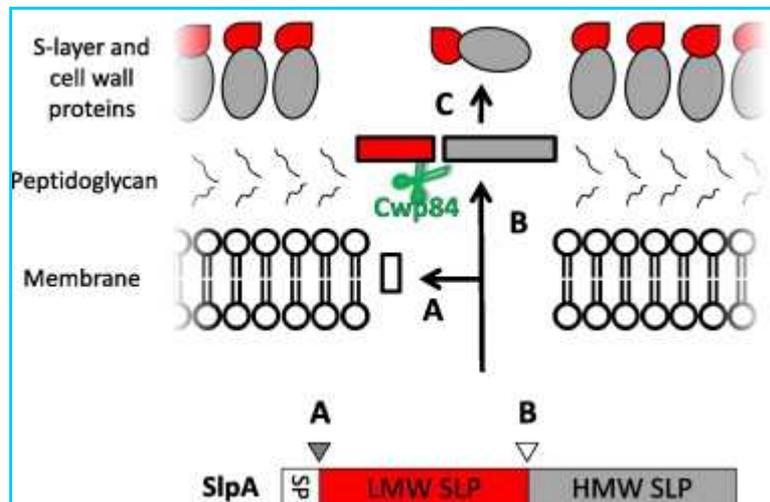


figure1. 5 Post-translational processing of SlpA.

The high- and low-molecular weight (HMW and LMW) surface layer proteins (SLPs) are synthesized in *C. difficile* as a ‘full-length’ precursor, SlpA. (A) During the process of translocation across the membrane the signal peptide (SP) is removed, presumably by a signal peptidase; (B) further processing by the papain class cysteine protease Cwp84 generates the mature S-layer proteins; (C) LMW and HMW SLPs self-assemble to form the S-layer. Taken from *Bioorganic & Medicinal Chemistry* 20 (2012) 614–621.

1.10.4 CwpV

This is another surface-exposed protein of *C. difficile* but one which has phase variable expression. Analysis has shown that it possesses three PF04122 cell wall binding domains and nine repeats of sequences each of 120 amino acids. The number of these repeat sequences appeared to differ between *C. difficile* strains. Domain-specific antibodies revealed that post translational modification takes place to cleave the protein, that the carboxy terminal part of CwpV is exposed on the bacterial surface but that at any one time, only 5% of bacteria expressed the

protein. Inverted repeats flanking the *cwpV* promoter and the action of a recombinase appear to explain the phase variable expression of *cwpV*. The exact role of CwpV is not clear but investigators suggest that it might help escape the host immune system or cycles of binding and release from host cells (Emerson et al., 2009).

1.10.5 Flagellum

In bacteria, the flagellum is made of three parts: the basal body, the hook and the filament. In *C. difficile* 630 the regions of the genome that are involved in flagellar synthesis are called loci F1 (CD0226-CD0240) and F3 (CD0245-CD0271) which are separated by a locus (F2) unrelated to this aspect of the biology of the organism (Stabler et al., 2009). Two structural components of the flagellum are FliC (the flagellin) and FliD (the flagellar cap protein). Preliminary studies of *C. difficile* strains showed the presence of a 39 kDa flagellar protein in bacteria of serogroups A (A1-A12), D, G, H and K (Delmee et al., 1990). Analysis of a 1.6 kb fragment from a genomic library revealed two orfs, one of 870 nucleotides and a second lying 90 bases downstream. Sequencing of *orf1* showed an amino acid sequence very similar to that of other flagellins – this was designated *fliC* of *C. difficile*.

The predicted molecular weight of FliC (30.9 kDa) was substantially less than that of the protein as determined by SDS-PAGE (39 kDa) suggesting post translational modification. Restriction analysis of *fliC* from six strains of *C. difficile* allowed separation into two groups. This can be explained by a region of high variability. Using electron microscopy, it has been found that some strains of *C. difficile* possess numerous flagella (eg strains W1194 and Kohn) but others carry fewer. Inhibition assays using anti-FliC sera showed that the flagellum is not able to adhere to eukaryotic cells (Tasteyre et al., 2000a).

Further study (Tasteyre et al., 2001b) confirmed the presence of a single copy of *fliC* with 46 strains possessing a gene of 870 bp. Organisms of serogroup X have a slightly shorter sequence (850 bp). *fliC* has also been found in 17 non-flagellated

strains and appears to be transcribed (RT-PCR) suggesting the absence of flagellae arises from mutation(s) in other genes. As found in previous work, amino acid alignment showed substantial variability in a central region with conservation at the amino and carboxy termini. Based on RFLP analysis, 47 strains could be divided into nine groups with most strains belonging to groups I and VII. RFLP groups II, III, IV, V, VI and IX all belong to a single serogroup but the rest are serologically distinct (Tasteyre et al., 2000b).

The flagellum filament is capped by FliD. Using primers for amplification of *fliD* gene from strain 630, a product of consistent size (1524 bp) was detected from 43 strains of *C. difficile*. The experimental and predicted molecular mass of FliD are the same (56 kDa). Data from RT-PCR with non-flagellated strains demonstrated again that this gene is transcribed in all of these strains. Moreover, FliD is highly conserved between strains and its surface exposure may contribute to attachment of bacteria to host cells (Tasteyre et al., 2001b).

Recombinant FliC, FliD and crude flagella have all been used to assess their role in adherence and colonization *in vitro* and *in vivo*. Mucus preparation from germ free mice was used in a blotting assay and the results showed recombinant FliC, FliD and crude flagella could bind. Colonisation of the mouse caecum was found to be more efficient for flagellated over non-flagellated strains. It therefore seems possible that flagellar proteins, particularly FliD, may contribute to adherence (Tasteyre et al., 2001a). In patients, both FliC and FliD have been shown to be immunogenic (Pechine et al., 2005b) and the vaccine potential of FliD has been tested in mice using different immunisation routes (intranasal, intragastric and rectal) and in combination with other surface proteins from *C. difficile* (Cwp84, Cwp66 and SlpA).

Although some effects on colonisation were apparent, it was concluded that more studies were required if an effective vaccine for *C. difficile* was to be developed (Pechine et al., 2007). Glycosylation appears important in assembly of

the bacterial flagellum (Logan, 2006) and a glycosyltransferase (CD0240) has been identified downstream of *fliC* in the *C. difficile* 630 genome. The predicted molecular mass of FliC from different strains is consistently less than that observed experimentally as noted earlier. In FliC from strain 630, five sites have been detected at S141, S174, T183, S188 and S205 at which O-linked glycosylation takes place; this explains the 9 kDa difference between predicted and observed molecular weights. Inactivation of the glycosyltransferase gene (CD0240) generated mutants without motile behaviour (Twine et al., 2009).

1.10.6 *Fbp68*

Fibronectin is one of the glycoproteins of the extracellular matrix with the ability to bind to integrin, collagen, fibrin and heparin sulphate proteoglycans. A gene was identified in *C. difficile* 79-685 for a 68 kDa fibronectin binding protein. Identification was made using the gene for a putative fibronectin binding protein from *B. subtilis* (GenBank accession number G69877). The gene is termed *fbpA* in strain 630. The gene (1773 bp) encodes a protein with 591 amino acids that are mostly hydrophilic in character that creates a two domain protein with different repeats within each domain, and a coiled coil structure in the middle of the protein. Fbp68 is highly conserved between strains 79-685 and 630.

Results of binding assays revealed that Fbp68 could bind both soluble and immobilised forms of fibronectin suggesting a mechanism for *C. difficile* to attach indirectly to cell surfaces (Hennequin et al., 2003). Although binding activity showed dependence on a number of metal ions, manganese appeared the most important in different aspects of Fbp68 interactions (Lin et al.). Binding of five manganese atoms to Fbp68 was shown to induce conformational change mainly in the random coiled region and to influence resistance to heat and protease activity. Interaction between Fbp68 and fibronectin appeared to be mediated *via* the carboxy terminal region of Fbp68 (residues 397-591) and the amino terminal

domain of fibronectin. Fbp68 is exposed at the surface of *C. difficile*. The role of the protein in adherence to mammalian cells *via* fibronectin was explored in experiments with *fbp68* mutants generated using Clostron technology. These experiments question the role of FbpA in colonisation (Barketi-Klai et al.).

1.10.7 GroEL (Hsp60)

This protein belongs to a large family of chaperones which are responsible for maintaining correct folding in their substrates. In a study of the properties of heated-shocked *C. difficile*, the rate of binding of bacteria to cells such as Caco-2 and HT29-MTX was increased (Eveillard et al., 1993). DNA analysis of strain 79-685 showed a single copy of a gene of 1940 bp encoding for GroES and GroEL (10 kDa and 57.6 K Da respectively). Upstream of *groES* a CIRCE-like sequence (controlling inverted repeat of chaperone expression) was detected as a binding site for HcrA, a negative regulator of heat shock protein expression. PCR-RFLP analysis of *groEL* from 12 strains indicated conservation in *C. difficile*.

Analysis of different fractions from heat-treated (48°C) *C. difficile* using antisera against recombinant GroEL showed this protein could be found in all fractions, but that it appeared mainly cytoplasmic and membrane fractions. Further localisation experiments showed that GroEL levels increased at the bacterial surface and appeared in the supernatant following heat shock. Expression was also enhanced by contact with eukaryotic cells and antisera decreased the binding of *C. difficile* to target cells in tissue culture (Hennequin et al., 2001b). Previous studies showed the importance of the local environment on the growth of *C. difficile* and the expression of different proteins (Waligora et al., 1999). With this in mind, the expression of GroEL was assessed under different stress conditions (heat shock, decreasing iron, acidic medium, high osmolarity and semi minimal inhibitory concentrations of antibiotic). It was found that stress has direct effect on GroEL expression (Hennequin et al., 2001a).

1.10.8 Peptidoglycan hydrolase

Peptidoglycan hydrolases or autolysins are found in different bacteria and have different physiological roles, for instance in cell division, growth or permeabilisation of the cell wall to allow passage of large components and extracellular assembly (Vollmer et al., 2008). A gene for a putative peptidoglycan hydrolase was identified in *C. difficile* 630 upstream of the gene encoding DNA polymerase III. The gene (1824 bp) was named *acd* and encoded a protein of 65.8 kDa protein with 607 amino acids and a pI of 9.52. Alignment showed the first 22 amino acids comprised the signal peptide, four repeat sequences similar to the bacterial SH3 domain were identified at the amino terminus of the mature protein and a catalytic domain with similarity to glucosaminidase enzymes was located at the carboxy terminus. The *acd* gene was conserved in sequence among 12 strains of *C. difficile* and was transcribed largely during the vegetative phase of growth, suggesting an important physiological role (Dhalluin et al., 2005).

1.10.9 Other surface proteins

Preceding sections only describe a selection of proteins thought to be present at the surface of *C. difficile*. A comparative genomic study among some members of the genus *Clostridium* revealed several domains that may determine surface location. These include the following Pfams: PF04122, a cell wall binding domain present in many surface proteins; PF01473, a domain that can anchor proteins at the surface through recognition of choline; PF00037 that binds to iron-sulfur; LPxTG, a motif that often acts as a substrate for sortase enzymes; PF04203, the sortase family itself; PF01471, a putative peptidoglycan binding domain; PF00232, an O-glycosyl hydrolase (Bruggemann and Gottschalk, 2008).

The sortases have been identified in gram positive bacteria as enzymes that can recognise exported proteins carrying appropriate motifs and couple them to peptidoglycan through a transpeptidation reaction (Paterson and Mitchell, 2004).

In *C. difficile* 630 only one putative sortase has been identified (CD2718) along with 14 putative sortase substrates, four of which appear to lie upon conjugative transposons. Of the putative substrates, two are collagen binding surface proteins (CD3392 and CD0386) a further two (CD0420 and CD1858) have been identified as putative cell surface proteins (Sebahia et al., 2006). In addition, analysis of the genome showed there to be proteases that may be membrane associated (CD0181, CD2129, CD1751 and CD0156) and a cell wall hydrolase (CD0183). These examples show the potential range of factors that may exist at the bacterial surface and may thereby contribute to the host-pathogen interaction.

1.11 Surface proteins and the aims of this project

In this project, it is proposed that a panel of recombinant antibodies be generated against some of the candidate surface exposed proteins highlighted above. The specificity of these reagents, the ability to express and purify them quickly and conveniently and to detect their binding and location using antibodies against small tag sequences would allow better definition of the location of the clostridial proteins, and conditions under which they are expressed. In addition, a panel of small recombinant antibodies could be tested to assess whether they show the ability to inhibit the natural activity of their targets and hence their individual contributions to the behaviour of *C. difficile* and its pathogenesis. Targets to be used comprise HMW and LMW derivatives of SlpA, Cwp66, Cwp84, FliC, FliD, and the putative sortase, FbpA, GroEL and CspA.

These will be expressed and purified from *E. coli*. To select the recombinant antibodies, a humanised phage library will be used to select high affinity binders against the recombinant targets. The selected antibodies will then be tested for binding to the recombinant proteins. Recognition of the native targets from *C. difficile* will also be assessed. Strong binders will be used to check if these antibodies can locate their targets at the surface of *C. difficile*. As a generic

approach to the analysis of location and activity of target proteins, these methods could in future be applied to proteins that make contributions to the biology of the pathogen.

1.12 Phage display

1.12.1 Principles of phage display

Phage display is a powerful technology that can be applied in many area of biological research. Basically, bacteriophage is used to display candidate peptides or proteins that are fused to components of the phage particle. The coding sequence for the peptide or protein is carried within the phage either as an insert to the viral genome or in a phagemid. Phage are isolated from mixtures by the ability of the peptide or protein to bind to a target. Therefore there is a direct link between phenotype (ability to interact with target) and genotype (coding sequence) which allows large libraries to be screened in a selection procedure. The first developments in this field were reported in 1985 by George Smith who confirmed that bacteriophage M13 could be used to present a foreign protein at the phage surface. To do this, a fusion protein was prepared by insertion of a DNA fragment into gene III of phage M13. The displayed protein retained native conformation and the infectivity of the recombinant phage was preserved allowing infection into suitable strains of *E. coli* and replication. Hence small numbers of phage with the capacity to bind could be amplified (Smith, 1985b).

The development of display libraries was reported shortly after Smith's initial paper. Amongst the first did an epitope library comprise millions of phage that displayed randomised peptides fused to pIII, the minor coat protein encoded by gene III. The library could used to separate those phage able to bind specifically to antibodies and hence map the epitopes recognised by these immunoglobulins

(Scott and Smith, 1990). Devlin reported a random 15-mer peptide library on M13 phage. Again, the peptides were fused into protein pIII of phage creating a library with a theoretical capacity of 2×10^{15} different phage. This library was used to identify peptides which could bind to protein targets (Devlin et al., 1990). Another early library was constructed by expressing random hexapeptide sequences at the amino terminus of adsorption protein of fd phage (Cwirla et al., 1990).

Filamentous phages (Ff) are amongst the most popular vehicles for phage display. They have a long cylindrical protein capsid which encloses a single stranded DNA genome. The genome was completely sequenced in the early 1980s (Hill and Petersen, 1982, Vanwezenbeek et al., 1980). The designation Ff is derived from their shape and their requirement for the F pilus for infection into *E. coli*. Their life cycle comprises infection, gene expression and replication, an assembly reaction and release from host by extrusion rather than lysis. Filamentous phages that have been used for phage display include M13, f1, Fd and ft. The dimensions of Fd phage particles are 7 nm diameter and 930 nm length. Their genomes are 6400 nucleotides in length and code for 11 proteins that contribute to assembly (pI, IV and pXI), replication (pII and X) or of the structure of the virion (pIII, pVI, pVII, pVIII and pIX).

The capsid of M13 virus is made up of 2700 copies of protein pVIII, a protein of 50 amino acids which can embed itself into membrane of the bacterial host prior to viral assembly through hydrophobic residues at positions 20-40. The amino terminus projects into the host periplasm while the carboxy terminus remains in the cytoplasm (Wickner, 1976, Ohkawa and Webster, 1981, McDonnell et al., 1993). These proteins coat the assembling virion as it is extruded through the bacterial membrane. Five copies of the pIII protein of M13 cap the particle on release from the bacterial host. Whilst there are reports of the use of pVIII for phage display, the low copy number of pIII has advantages for display technology (see below) (Arap, 2005).

pIII (406 amino acids) carries an 18 residue amino terminal signal sequence for export *via* the bacterial Sec protein system (Rapoza and Webster, 1993). After cleavage of the signal peptide, folding of the protein to its mature form creates two amino terminal domains (N1 and N2) and the CT domain at the carboxy terminal end which all are linked to each other by glycine rich sequences. The CT domain (residues 256-406) of pIII is partially buried in the tip of phage particle. This holds pIII in place and the interaction ends the process of phage assembly and initiates releasing of the extruded virion.

Approach to a new bacterial host commences a new cycle of infection, a process to which pIII also contributes. Analysis of crystal structures has shown that the binding of the N2 domain (residues 87-217) to the bacterial F pilus induces the release of the N1 domain (residues 1-67) from a intramolecular interaction. This then allows N1 to bind TolA in the bacterial outer membrane; TolA thus acts as co receptor (Holliger et al., 1999, Russel et al., 1988, Rakonjac et al., 1999). Further studies have suggested three step model for attachment with initial interaction of N2 with the F pilus, binding of N1 to TolA and in a subsequent step, attachment of N2 to a central domain of TolA (Karlsson et al., 2003). These intricate interactions can be retained after attachment of peptides and proteins to pIII or through the use of helper phage, loss of function can be compensated.

Whilst peptides and proteins can be fused with gIII in the viral genome, phagemids provide a more convenient basis for phage display. In a bacterial host, phagemids behave like conventional plasmids but they possess the ability to be packed into filamentous phage particles as single stranded DNA if the necessary functions are supplied *in trans*. This is possible because these vectors carry *ori* sequences from M13 phage and for initiation of plasmid replication. Helper phage such as M13KO7 provide the genes required for viral assemble and packaging but with a less effective origin of replication than that carried on the phagemid, the helper phage genome is incorporated into particles with much lower frequency than the phagemid.

Helper phages have been adapted to enhance these general properties. For example, Ex-phage was prepared by site-directed mutagenesis to create an amber codon in *gIII*. This helper phage can generate functional pIII only in suppressor strains of *E. coli* and has been used in antibody phage display applications (Baek et al., 2002). In another study, a helper phage known as CT-phage was produced in which domains N1 and N2 were deleted from pIII (Kramer et al., 2003). The helper phage used in the current study, KM10, was prepared with a protease cleavage site between the second and third domains of pIII and hence virus carrying copies of pIII encoded by KM10 (*ie* lacking additional sequence fused to pIII) are non-infective after protease treatment (Kristensen and Winter, 1998).

Different vector systems have been developed for expression of fused sequences to the capsid proteins of filamentous phage. As indicated earlier, pIII and pVIII have been most frequently exploited as structural studies of pVI, pVII and pIX proteins have failed to show promise (Makowski, 1993). In the first type of vector systems, two genes (either *gIII* or *gVIII*) are carried on separate genomes, the wild type gene carried on the helper phage and the recombinant version carrying fused sequence on the phagemid. These vectors are known as 3+3 or 8+8 systems. When bacteria carrying the recombinant phagemid are infected with a helper phage, virus carrying the phagemid genome are produced but the capsid carries a mixture of wild type and fusion protein (pIII or pVIII). The alternative systems are vectors that carry two copies of *gIII* or *gVIII* (so-called 33 or 88 type vectors). In these systems, the genome of phage encodes both a wild type and a recombinant version of the same sequence.

In either vector system, it is possible to regulate to an approximate extent the balance of wild type and recombinant capsid proteins. For pIII, low representation or monovalent display of the fusion protein has advantages as avidity effects in attachment of the phage to the target can be minimised and hence high affinity

binders can be extracted from libraries. For pVIII, the high copy number of this protein on each particle provides opportunities to select for low affinity interactions with target (Smith and Petrenko, 1997).

For both pIII and pVIII, foreign sequences are inserted after the signal peptide to ensure surface exposure at the surface of the phage particle (Smith, 1985a). Where library construction and fusion of additional sequence destroys the normal function of pIII, infectivity can be restored by protein encoded by the helper phage and the recombinant phage carry a mix of non-functional recombinant protein and function wild type (Felici et al., 1991, Greenwood et al., 1991). Vector systems developed for library construction have incorporated restriction sites for the insertion of foreign sequence, selectable markers (Scott and Smith, 1990, McCafferty et al., 1990a, Hoogenboom et al., 1991, McConnell et al., 1994) and epitope tags for detection of the fused sequence. One of the most frequently used tags has been the c-myc epitope (EQKLISEEDLN).

This can be found in vectors such as m663, M13mp18Xa enabling discrimination of recombinant and wild phage particles. In the pHEN display vector, the c-myc tag is located at the carboxy terminal end of the insert, before the coding sequence for pIII (Hoogenboom et al., 1991). Other modifications include the addition of a proteolytic cleavage site between foreign and capsid sequence (*eg* a site for factor Xa in M13mpXa vector (Rodriguez and Carrasco, 1995)) and *lacZ* (*eg* in vectors like m66) to enable the detection of inserted sequence by blue white screening with Xgal (Fowlkes et al., 1992). Stop codons between the displayed sequence and pIII enable the expression of a fusion with the capsid protein in suppressor strains of *E. coli* or expression of the foreign sequence in non-suppressors without modification of the vector (Hoogenboom et al., 1991, Lowman and Wells, 1993).

1.12.2 Type of libraries and application of phage display

Generally phage display libraries can be divided into those that carry randomised sequences (*eg* random peptide libraries) and those that display natural peptide or protein domains on the phage surface (*eg* cDNA libraries).

In this field, random peptide libraries (RPL) have developed quickly to support the identification of epitopes or synthetic mimotopes, as used in the early identification of random peptides against β -endrophin (Lam et al., 1993). The use of randomised synthetic oligonucleotides has enabled the convenient assembly of large libraries (*eg* NNK or NNS codons. N representing an equal mixture of all nucleotides; K, equal mixtures of G and T; S, equal mixtures of G and C). For instance a 6-mer RPL can have 64 million peptide sequences (Irving et al., 2001). Synthetic oligonucleotides generated in this way can be cloned in-frame at the amino terminal coding region of *gIII* or *gVIII* of M13. Once assembled, the RPL can be screened for clones with affinity against different targets such as the identification of immunodominant antigenic peptides, peptide competitors in antigen-antibody reactions or mapping functional sites. In this context, RPL have been used in HIV research (Christian et al., 1992), to characterise the hepatitis B coat protein (Folgori et al., 1994, Felici et al., 1993) and in studies of *Bordetella pertussis* toxin (Felici et al., 1993).

In research of receptor-ligand interactions, RPLs have proved equally valuable. Peptides from RPLs may possess similar primary sequence to receptor ligands or mimic the characteristics of peptide or non peptide ligands. In one application, platelets were used to screen an RPL to select peptide sequences able to bind to the thrombin receptor. Selected phage were able to precipitate the thrombin receptor and peptides were identified that behaved like tethered ligands (Doorbar and Winter, 1994). Antagonists against urokinase were also selected from a 15-mer RPL (Goodson et al., 1994), three peptides with high affinity for single stranded DNA were extracted from a phage library (Krook et al., 1994) and peptide mimics of a carbohydrate ligand were identified by phage display against

concanavalin A (Oldenburg et al., 1992). Other examples of the use of RPLs include studies to define substrate specificity. In this example, an RPL was prepared with an amino terminal sequence that mediated binding to a solid support. The RPL was bound to the affinity support and then treated with the proteases under investigation. Cleavage of sequences in the RPL then released the phage while virus carrying resistant sequences remained attached to the solid support. Sequencing of insert sequences carried by the phage revealed the substrate sequences (Matthews and Wells, 1993).

The other main use of phage display technology is to display libraries of proteins or their isolated domains and to select from these libraries, proteins with particular properties. Three areas of application are highlighted: the mutagenesis and directed evolution of protein domains; the screening of cDNA expression libraries; and in the next section, the isolation of high affinity antibodies.

Phage display can be used to direct the evolution of a protein or protein domain by creating multiple mutations, selecting for desirable properties, and repeating cycles of mutation and selection to create by increments particular properties within the protein. Different methods can be used for mutagenesis such as cassette mutagenesis, error prone PCR or DNA shuffling. By displaying the library of mutated protein sequences on phage, binders can be selected by attachment to a solid surface. This has been used in receptor studies (Lowman and Wells, 1993) and the development of novel enzyme inhibitors (Dennis and Lazarus, 1994). cDNA expression screening is other powerful technique. Here, proteins encoded in a cDNA library are displayed at the phage surface. This has been used to identify epitopes from antigens able to induce autoimmunity (Fierabracci et al., 1999). Another study used this method to target drugs for treatment of prostate cancer (Arap et al., 2002).

Similar studies of targeting have used *in vivo* selection; here, phage libraries are injected intravenously into animals, specific organs are isolated, and phage are

isolated that specifically accumulate because they carry proteins with a particular tropism. The first studies in this areas focused on the phage that could bind to the renal and cerebral vascular endothelium (Pasqualini and Ruoslahti, 1996) but the technique can be used to find antigens that are recognised by the antibodies circulating in patients with cancer (Mintz et al., 2003). Isolation and identification of allergens is another application of phage display. Phage display cDNA libraries can be screened for binding to serum IgE from patients. Phage isolated in this way can then be sequenced to identify the DNA sequence of the potential allergen (Rhyner et al., 2004).

While these applications focus upon the isolation and characterisation of reactive protein sequences, phage display can be also used in gene delivery to improve the efficacy of delivery and decrease toxicity (Goldman et al., 1997, LAROCCA et al., 1999, Barry et al., 1996) and, related, in development of DNA vaccines (Gao et al.) or conventional immunogens. In DNA vaccination, the phage can be used as a carrier of a DNA vaccine expression cassette in which the gene for the candidate antigen is under the control of an appropriate eukaryotic promoter (Clark and March, 2004a). Testing these ideas with the capsid antigen from hepatitis B virus showed that phage delivery could stimulate an effective and long lasting immune response in mice (Clark and March, 2004b). With more conventional approaches to vaccination, it has been shown that the display of immunogenic sequences on phage particles has advantages over the use of traditional protein carriers like ovalbumin revealed (van Houten et al., 2006). An orally delivered food and mouth disease vaccine was reported to induce 100% protection in mice (Ren et al., 2008).

1.12.3 *Antibody phage display*

1.12.3.1 A brief history and application

In the early 1890s, Emil Von Behring and Shibasaburo Kitasato revealed the sera of animals immune to diphtheria and tetanus contained an antitoxic activity which was much later attributed to the presence of circulating proteins that were termed antibodies. Antibodies are a group of protein molecules produced in response to foreign substances and thereby create an important defence against pathogens. An antibody can bind to specific part of a foreign molecule *via* interaction of its paratope with the corresponding epitope carried by the antigen. This reaction can be used in diagnosis, neutralisation of the antigen, localisation of the interaction or therapy. The specificity of the interaction and its affinity is determined by amino acids present in the complementary determining regions (CDRs) of the antibody. Of their nature, these are highly variable; other parts of the antibody – regions of the so-called variable domain and the constant domains of the heavy and light chains – are conserved to a lesser or greater extent.

The variable domains of heavy and light chains are encoded by gene segments that undergo rearrangement as VDJ (variable, diversity and join segments; heavy chain) or VJ (light chain) units (Litman et al., 1993, Borghesi and Milcarek, 2006). The multiplicity of responses that are stimulated by an antigen give rise to polyclonal responses which can be used to neutralise or detect toxic materials (Harris, 1999), for protection against pathogens (Casadevall, 2003), for research into disease states (Lambrianides and Giles), cancer therapy (Reichert and Valge-Archer, 2007) and many other branches of biology. Work in this area has been transformed by two landmark developments that took place in 1975 and 1988.

In 1975, it was shown that fusion of myeloma cells with B lymphocytes could be used to produce monoclonal antibodies against an antigen, thereby deconstructing the nature of the polyclonal response (Kohler and Milstein, 1975). Humanization of monoclonal antibodies was reported in 1988, a process in which the antigen binding properties of a rat antibody were transferred to the structure

of a human IgG (Riechmann et al., 1988). This helps overcome the immunogenicity of rodent antibodies in the human patient. This can be approached by the generation of chimeric antibodies that take the variable domains of a rodent antibody (and hence the capacity to recognise antigen with high specificity) and to fuse them with human constant domains. Antibodies of this type are 60-70% human in sequence and the limited response that they provoke in patients has led to their promotion for therapeutic use (Brekke and Sandlie, 2003, Chintalacharuvu and Morrison, 1995, Lupo et al., 2008).

The alternative approach is CDR-grafting that was developed in the 1990s. Here, CDRs from a specific rodent antibody are transplanted onto human frameworks creating a recombinant molecule that is 90-95% human antibody sequence (Jones et al., 1986, McCafferty and Glover, 2000). Literally applied, grafted antibodies failed to retain the affinity of the original immunoglobulin because the influence of residues from the framework was overlooked. Improvements can be achieved by choosing a recipient human sequence that shows high homology to that of the donor rodent molecule (Shearman et al., 1991).

A completely different approach is through the engineering of the mouse immune system such that antibodies of human sequence are expressed by mice that have been experimentally modified or fully transgenic animals ("HuMab mice"). Immunization thereby leads directly to the production of human antibodies that can be immortalised *via* hybridoma techniques (Green, 1999, Neuberger, 1996). Mice with these properties can be generated by transplanting human cells from for a human donor into animals that carry severe combined immunodeficiency or animals that have been lethally irradiated and then immunologically reconstituted (Reisner and Dagan, 1998). Mice that have been manipulated in this way can be also used as models for the study of different diseases such as hepatitis virus (Ilan et al., 1999). Transgenic approaches to the same challenge are also possible by creating mouse lines in which the endogenous

antibody responses is inactivated and replaced with a human mini chromosome containing the human heavy and kappa light chain loci (Tomizuka et al., 2000).

The structural basis of the interaction between antigen and antibody is now understood in substantial detail. The variable domain of each antibody heavy or light chain comprises a structure with two antiparallel beta sheets and an intramolecular disulfide bond. Three loops from each chain project from this structure and the amino acids in these short sections form the CDRs that together create a unique interface for the interaction with antigen. Four of the six loops – two in the light chain, two in the heavy chain – are initially encoded by germline segments, the others being created through the rearrangement process.

Substitutions can be incorporated at any position in the variable domains during somatic mutation as B cells respond to antigen, but substitution in the CDRs are particularly important for the nature of the interaction with antigen. While the third heavy chain CDR (H3) loop was found to be most variable in length and sequence, L3 loop (CDR3 in the light chain) was determined to form the central point area of the antigen binding site (Wu and Kabat, 1970). The antibody binding sites so created can be classified into 4 groups including concave and mild concave (these are found with particular frequency in the interaction with small antigen molecules), rigid (peptide antigens) and planar (protein binders) (MacCallum et al., 1996). Based on comparative studies, a series of “canonical” conformations have been identified that make possible the formation of these surfaces (Chothia et al., 1989). In addition, sequence analysis and comparative studies of different framework and CDR sequences have identified residues that are particularly critical. These include residue 71 which lies within a framework region but is particularly influential in the conformation of the second CDR in heavy chains (H2) (Tramontano et al., 1990).

Since 1975, scientists have worked towards the use of monoclonal antibodies in the clinic (Reichert, 2001). In attempts to overcome the reliance on

production of monoclonal reagents in cell culture, work through the 1980s began to look at cloning and expressing antibody genes in bacteria (Winter and Milstein, 1991). Amplification of the genes encoding monoclonal antibodies by PCR and expressing them in prokaryotic hosts like *E. coli* offers a convenient and economic route to antibody production (Plückthun et al., 1989). Whilst inherently variable in sequence, the rearrangement of segments from modest numbers of gene families has made this possible (Larrick et al., 1989) (Benhar and Pastan, 1994, Ruberti et al., 1994).

Expression of full-length antibodies in bacterial hosts is problematic but the proteins can be engineered into different formats such Fv fragments (VH and VL domains that spontaneously associate), scFv proteins (VH linked to VL by a peptide linker), Fab fragments (VH and VL domains with the first constant domains) and oligomeric units of two, three or four identical or distinct antibody fragments (“diabodies”, “triabodies” etc) (Figure 1.5). Fortuitously, the robust fold evident in the variable domain is retained in proteins with these formats and hence the capacity to interact with antigen and the affinity of the interaction is retained.

Amongst these engineered antibodies, the most commonly encountered formats are scFvs and Fabs. In scFvs a short polypeptide linker can be fused to link VH and VL allowing the domains to fold independently and associate in the bacterial periplasm but retaining them in close proximity as a single protein chain (Winter et al., 1994). Fabs contain two polypeptide chains that are independently coded and translated. The two chains can interact through formation of a disulfide bond which increases the stability (Carter et al., 1992) (Figure 1.5).

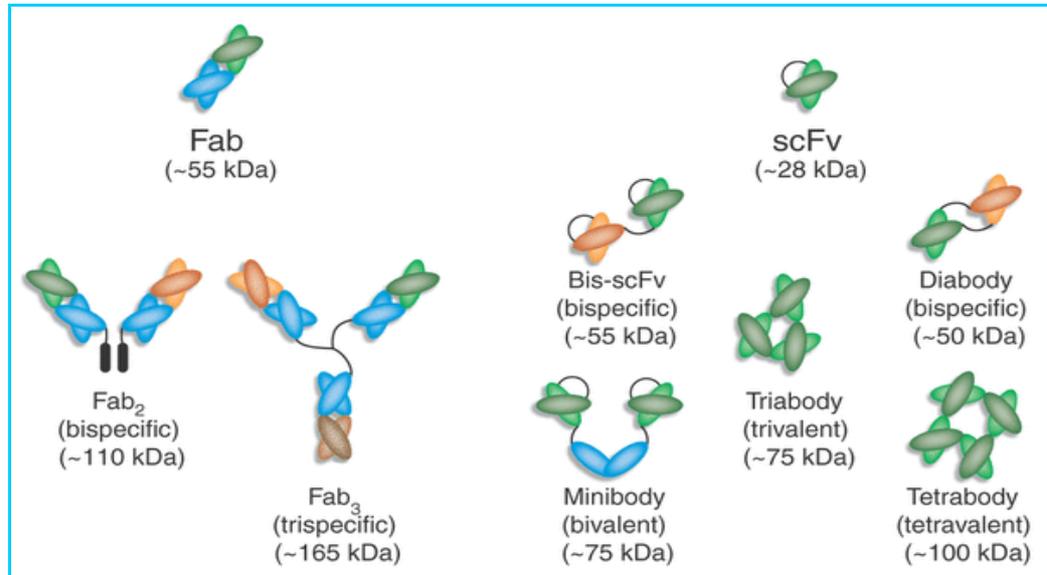


figure1. 6. Formats of engineered antibodies (taken from Nature Biotechnology 29, 5–6, [2011]).

1.12.3.2 Formats of antibodies in phage display

1.12.3.2.1 Preface

Antibody phage display is identical in its approach to the other formats of phage technology described earlier: antibody chains are expressed as fusion proteins with one of the components of the phage coat; the recombinant phage particles are presented with target that is coated to a solid surface and those that fail to interact are washed out; phage recovered from the selecting surface are infected into bacteria and replicated in preparation for a further round of selection. As a technology, antibody display benefited from the coming together of display methods (McCafferty et al., 1990b) and the appreciation that export of antibodies to the environment of the Gram negative periplasmic would allow folding to take place, thus creating functional immunoglobulin-like proteins (Skerra and Pluckthun, 1988b).

Recombinant antibody sequences can be carried on phage vectors but the use of phagemids vectors dominates the field. In these systems, the antibody sequence (typically scFv) is cloned into *gIII* in frame with the coding sequence for the minor capsid component described earlier (Clackson et al., 1991, Hoogenboom et al., 1991, Barbas et al., 1991). Helper phage such as KM13 or M13KO7 are used to catalyse the assembly and packaging of phagemid into particles that display the antibody-pIII fusion (Kristensen and Winter, 1998, Vieira et al., 1987). The most efficient antibody formats for display are scFv (Marks et al., 1991), Fabs (Chang et al., 1991), Fvs and diabodies (Brinkmann et al., 1995, McGuinness et al., 1996) (Figure 1.5). As regards the range of antigen-binding capacities contained in a library of antibodies, these are made available through four approaches.

1.12.3.2.2 Immune libraries

Basically this method samples the antibody repertoire of an individual that has been exposed to an antigen (*eg* a pathogen) or deliberately immunised, using specific primers to recover antibody sequences by PCR (Welschof et al., 1997). The composition of the library will thus be influenced by the antigen(s) used in immunisation of the human patient or experimental animal (Cai and Garen, 1995, Barbas et al., 1993) but libraries of this sort are likely to be enriched for specific antibodies with high binding affinities (Clackson et al., 1991). Immune libraries have been reported from different species including humans (Barbas et al., 1993), mice (Kettleborough et al., 1994), chickens (Davies et al., 1995), rabbits (Yamanaka et al., 1996), and camels (Arbabi Ghahroudi et al., 1997) amongst others.

1.12.3.2.3 Naïve libraries

The term “naïve” has been applied to these resources as they are prepared from non-immune sources that have not been stimulated to respond towards a specific target. Providing the libraries are large (approximately 10^7 different clones or greater) antibodies with different binding affinities can be found against a great many antigens in libraries of this type. Using PCR, recovery of sequences can draw upon IgM antibodies present on the surface of non activated B cells.

During development of this area, investigators have sought to create very large libraries to increase the chances of isolating high affinity binders to the target of interest (Marks et al., 1991, Griffiths et al., 1993). Affinity ranges for a small library of around 3×10^7 clones lie between 10^6 - 10^7 M^{-1} whereas creating much larger libraries (10^{10} clones) has been shown to yield antibodies with affinities in the range 10^8 - 10^{10} M^{-1} (de Wildt et al., 1996, Griffiths et al., 1993, Vaughan et al., 1996). Two or three cloning steps are used to create libraries of this size. Firstly the

amplified sequences of light chains are cloned to the phage vector and in a second step, heavy chain amplicons are ligated (Welschof et al., 1997, Johansen et al., 1995). Some authors have divided the second step such that the heavy chain repertoire is constructed in a holding vector before insertion into the light library (Welschof et al., 1997). The random pairing of light and heavy chain products enhances the diversity of the resource.

1.12.3.2.4 Synthetic antibody libraries

In this type of antibody library, variable genes are assembled together and diversity is created in the CDRs and / or bordering framework using *in vitro* methods (Hoogenboom and Winter, 1992, Barbas et al., 1992). The range of diversification and its location is chosen from studies of the natural diversity of antibody from *in vivo* sources, often focussing on CDR3 of the heavy chain (Chothia et al., 1989). In construction of synthetic libraries, there are many reports that used a variety of sequences and approaches to libraries containing large numbers of antibodies with reasonable efficiency.

Some of the creators have used as scaffolds the frameworks of known antibodies and incorporated randomised CDR3 sequences to the heavy and light variable chains regions (Desiderio et al., 2001). Another technique for constructing synthetic libraries was to use the diversity created *in vivo*. In this approach, CDRs were isolated from *in vivo* sources and integrated into a single human framework (Jirholt et al., 1998). Fully synthetic library was also reported that used heavy and light frameworks and incorporated six synthetic CDR cassettes (Knappik et al., 2000). One of the largest synthetic libraries was constructed by combining of 49 heavy human heavy chain sequences with a collection of 47 human κ and λ light chains, each containing partially randomized CDR3 sequences. Antibodies against different antigens were selected from this library with nanomolar affinities towards their targets (Griffiths et al., 1994).

One advantage that has been exploited in the development of synthetic libraries is that variable gene segments can be chosen that are well-expressed in the bacterial host, that fold successfully and that possess low toxicity in bacteria. These features enhance the functional size of the library. In contrast, one problem that arises from *in vitro* diversification – particularly when sequences are randomised – is the creation of stop codons. This can be addressed by assembling from trinucleotides instead of single bases (Virnekas et al., 1994). Further design considerations can include pre-selection of synthetic variable domains on immunoglobulin binding proteins such as Protein A (for VH) and Protein L (for VK) (Ikerström et al., 1994), ensuring that all components of the library are functional.

1.12.3.3 Selection strategy

Having created a phage display library of diversified antibody sequences in scFv, Fab or other formats, phage that are reactive with a target can be extracted by panning. In panning, the target is immobilised to a solid support (*eg* a plastic surface with the capacity to bind protein) and the phage library is added. During washing, phage that lack the capacity to interact with target are eliminated and those that are captured to the solid support are recovered, infected into a bacterial host to allow replication from initial small numbers to yields sufficient for the next round of selection. Repeated rounds of selection, recovery and replication progressively enrich for binders.

There many different selection methods, the simplest being biopanning on supports to which the target “antigen” is immobilised (Clackson et al., 1991), be that a tube, columns (Marks et al., 1991, McCafferty et al., 1990b), or BIAcore sensor chips (Malmberg et al., 1996). In some cases, purification of the intended target proteins may lead to loss of the biological activities required for selection. Under these circumstances panning is possible on prokaryotic or eukaryotic cells (Bradbury et al., 1993, Cai and Garen, 1995) or in living animals (Pasqualini and Ruoslahti, 1996); the principles are identical but libraries may need to be pre-

absorbed with cells that lack the target to ensure selection is guided specifically towards the target and spurious features of the cell surface are avoided (Watkins and Ouwehand, 2000).

Given that very large numbers of phage are added to the selecting surface, the frequency and stringency of washing is an important factor in ensuring that panning is efficient and specific. Limited numbers of washing cycles are used to isolate low affinity antibodies while high affinity antibodies can be extracted from libraries if more stringent conditions are employed (Winter et al., 1994). Other interventions that can drive the isolation of high affinity binders include use of low concentrations of target at the selecting surface. The number of rounds of selection required for extraction of specific antibodies from the display library depends upon the enrichment rate at each stage but three rounds of selection are typically employed given that enrichment rates are in the range of 20 to 1000 fold at each selective step (Marks et al., 1991, McCafferty et al., 1990b).

Recovery of those phage – often very small numbers of phage from the library – from the target-coated surface can be through the use of media of low pH such as glycine buffers (Kang et al., 1991, Roberts et al., 1992) or high pH such as triethylamine (Marks et al., 1991). Chaotropic agents with dithiothreitol have been reported; these are usual when biotin is linked to the target molecule through a disulfide bond (Griffiths et al., 1993). Enzymatic cleavage is also a possibility that can be usefully exploited when cleavage sites have been engineered in place between the antibody and pIII (Ward et al., 1996). Competing concentrations of free target can also be used (Clackson et al., 1991).

1.12.3.4 Phage antibody screening

In the selection steps of phage display, combinations of binders with different properties are recovered from the initial library. To identify clones that are best-suited to the intended end use, screening methods are used that are fast, robust, and able to work with unpurified phage or soluble antibodies that are derived from

the phage. Mixed (“polyclonal”) populations of phage or clonal lines (“monoclonal”) phage antibodies can be tested in ELISA to assess the specificity of their interaction with target (Marks et al., 1991). Immunoprecipitation (de Wildt et al., 1996) or immunocyto chemistry (VanEwijk et al., 1997) can also be used to confirm specificity. Bioassay screens that test directly if phage antibodies can neutralise the biological activity of the target are also possible (Zaccolo et al., 1997).

One value of modern phagemid vectors is that many include an amber codon between the antibody coding sequence and *gIII*, allowing the screening of antibodies at the phage surface and the convenient production of soluble antibodies as monovalent proteins by switching from suppressor to non-suppressor strains of *E. coli* (Hoogenboom et al., 1991). The incorporation of epitope tags to these vectors allows phage or protein to be detected immunochemically with reagents against c-myc (Marks et al., 1991) or Flag sequences (Lah et al., 1994). The other tag of value that these vector often fuse to the antibody sequence is a His tag that enable convenient purification of soluble antibody *via* affinity chromatography (Hochuli et al., 1988).

1.12.3.5 Antibody expression

As outlined earlier, antibody phage display is only possible because microorganisms such as *E. coli* can be used as hosts for expression of recombinant antibodies. Once clones emerge from phage display, selection and screening, their properties can be best characterised through large scale expression and purification.

Antibody Fab proteins contain five disulfide bonds, and scFvs contain two; successful folding and the formation of these bonds takes place in the Gram negative periplasm. To attain this location phage display vectors provide bacterial leader peptides derived from *pelB*, *phoA* or *ompA* (Skerra, 1993, Skerra and Pluckthun, 1988a). Within the periplasmic environment, bacteria proteins like

DcbA, DsbB and DsbC assist folding and the formation of disulfide bonds. Yields of proteins from the bacterial periplasm can be enhanced by the use of specialised expression vectors with strong promoters (Colcher et al., 1990, Gibbs et al., 1991). Protein misfolding and aggregation can be decreased by growing cells at 25°C or applying tightly regulated promoters such as pBAD arabinose inducible promoter (Somerville et al., 1994, Clark et al., 1997).

Some investigators have reported the use of non-metabolised materials to create osmotic stress and thereby increase the yield of correctly folded antibodies in the periplasm (Kipriyanov et al., 1997, Kipriyanov and Little, 1999). Codon usage also appears to have effects in some cases on the solubility and hence the yield of antibodies from prokaryotic expression hosts (Dueñas et al., 1995, Forsberg et al., 1997) and as described earlier, grafting CDRs onto the framework of antibodies that are inherently well-expressed also improves stability and antibody production (Würl and Plückthun, 1999).

1.13 Aims of the project

Bringing together the themes described above, this project aimed to generate a panel of recombinant antibodies against proteins known or suspected to be present at the surface of *C. difficile* the goals being to assess surface exposure where this is uncertain, to test whether the function of these proteins can be blocked by specific recombinant antibodies, and thereby to generate reagents that may contribute to a better understanding of the role of surface proteins in the biology and virulence of this important nosocomial pathogen.

2 Material and methods

2.1 Culture of *Clostridium difficile*

To prepare genomic DNA, flagella and extracts containing surface layer proteins, *C. difficile* was cultured in liquid and agar media. Three strains were used at points in the project: 630 (the first strain to be sequenced; GCC 107), R20291 (hypertoxigenic outbreak strain, ribotype 027; GCC 154) and M120 (a non-motile strain of ribotype 078; GCC 165). These strains were obtained as spores from the Glasgow Culture Collection maintained by Dr G. Douce (University of Glasgow) and grown under anaerobic conditions at 37°C.

Brazier's CCEY agar and BHI (brain heart infusion) were used to culture the strains. For the former medium, 48 g of Brazier's powder (Oxoid) was added to 1 l deionised water, mixed and autoclaved at 121°C for 15 minutes. The autoclaved medium was allowed to cool to 47°C and then 10 ml cycloserine/cefoxitin with 40 ml egg yolk emulsion was added as a supplement (X093 and X073, LAB M). The supplemented medium was mixed well and poured into petri dishes. Plates were put on the bench to solidify and then kept at 4°C until required. BHI (Oxoid) liquid medium was prepared by weighing 12 g of powder into 1 l deionised water which was then autoclaved at 121°C for 15 minutes.

C. difficile was initially plated on Brazier's CCEY media and incubated for 48 hours under anaerobic conditions at 37°C. After 48 hours, the plates were checked for growth, the appearance of colonies with a ruffled morphology and the smell that is characteristic of *C. difficile*. Aliquots of BHI liquid medium were pre-incubated under anaerobic conditions overnight with caps loosened to allow removal of oxygen. Cultures were inoculated with single colonies from agar plates and grown under anaerobic conditions at 37°C.

2.2 Selection of target proteins

Surface proteins of bacterial pathogens operate at the interface with the host and make important contributions such as chemotaxis towards host cells, attachment and pathological change to host tissues. Whilst detailed structures are limited in their availability, analysis of hydrophobicity plots also can be a useful tool in such features as transmembrane domains or regions of the protein that are likely to be exposed. Kyte-Doolittle plots are commonly used to analyse the hydrophobicity of a protein while the Hopp-Woods scale can predict hydrophilic domains. Use of these algorithms and a survey of literature was used to select the following for cloning and expression as target proteins: SlpA, Cwp84, Cwp66, Acd, GroEL, FliD, FliC, CspA, FbpA and a putative sortase, Sortase B.

2.3 Genomic DNA extraction

Genomic DNA from strain 630 was used as template to amplify coding sequences for the selected targets. DNA was extracted using a genomic DNA purification kit from Promega. To do this, a 10 ml BHI liquid culture of the bacteria was set up by inoculating a single colony of strain and incubating in an anaerobic cabinet at 37°C overnight. The overnight culture was dispensed in 1 ml aliquots to 1.5 ml tubes and centrifuged at 12000 x g for 2 minutes in a bench top centrifuge (Thermo Scientific PICO17). After removing the supernatant, the pellet from each tube was resuspended in 480 µl of 50 mM EDTA and 120 µl of 10 mg/ml lysozyme (Sigma) was added to the resuspended cells.

This was mixed gently and incubated at 37°C for 30-60 minutes before spinning again for 2 minutes at 12000 x g to remove the supernatant. The remaining extraction process was carried out according to the technical manual: in turn, this lysed the cells, degraded RNA by addition of RNase, precipitated the cellular protein, and finally recovered genomic DNA by precipitation in 70% (v/v). The

amount of DNA from each tube was measured using a Nanodrop instrument and samples with high yields were mixed and aliquoted in 50 µl samples for storage at -20°C.

2.4 Primer design and PCR

2.4.1 Primer design

The genome sequence of *Clostridium difficile* 630 (NC-009089.1) was used to design all primers for amplification of the selected coding sequences. For each target, sense (forward) and antisense (reverse) primers were designed with additional sequence at the 5' terminus to enable ligation independent cloning (LIC) into the pET32 EK/LIC vector (Novagen). According to the LIC system, the 5' terminus of the coding sequence requires a 5' overhanging sequence of **5' GAC GAC GAC AAG ATN 3'** followed by the insert-specific sequence. **N** can be any base but this then forms a codon that determines the reading frame in which the expressed sequence is translated. The 3' terminus of the coding sequence requires a 5' overhanging sequence **5' GAG GAG AAG CCC GGT 3'** followed by the intended insert sequence. Table 2.1 shows the sense and antisense primers that were designed for each target and used in PCR to amplify the chosen coding sequence.

Table 2. 1. List of sense and antisense primers

GENE	SENSE PRIMER	ANTISENSE PRIMER
groEL	5' GAC GAC GAC AAG ATT GGA GTA ACT ATA GCA AAA GAG 3'	5' GA GGA GAA GCC CGG TCC GCC ACC CAT TCC TGG 3'
cwp84	5' GAC GAC GAC AAG ATA GAT GGA GTA GAA ACT GCA GAG 3'	5' GA GGA GAA GCC CGG TTC ATT TCC ATT TCC ACC AAC 3'
cspA	5' GAC GAC GAC AAG ATG AAA AAC GGA ATA GTA AAA TGG 3'	5' GA GGA GAA GCC CGG TAC GTT TTC AGC TTG AGG TCC 3'
5 cwp66	5' GAC GAC GAC AAG ATA ACG GGT TCT GGA AGA TGG 3'	5' GA GGA GAA GCC CGG TTT AGC TGC TAA TAC ACC CAC 3'
3 cwp66	5' GAC GAC GAC AAG ATA GTT ACT CAA ATT GGT GGC 3'	5' GA GGA GAA GCC CGG TGT AGC GTT AGT TCT TCC 3'
fbpA	5' GAC GAC GAC AAG ATA CAT CAA CCT GAA GAT GAT GAG 3'	5' GA GGA GAA GCC CGG TTT AAC CTT AAG CTT GGC TAC 3'
fliC	5' GAC GAC GAC AAG ATG GAG AAG TTA TCT TCT GGG 3'	5' GA GGA GAA GCC CGG TAA AAC TCC TTG TGG TTG TTG 3'
fliD	5' GAC GAC GAC AAG ATT CCA GTA AGA GTT ACA GGC 3'	5' GA GGA GAA GCC CGG TTG TGA GAA ATA GTT CAT TTG 3'
Acd	5' GAC GAC GAC AAG ATT GAA CCA ACT GCC GAA AGT AGC 3'	5' GA GGA GAA GCC CGG TTC CAT AAT TCC AGA AAT TCC 3'
sortaseB	5' GAC GAC GAC AAG ATC AAT CAT GAT ACT AAA ATA TCC 3'	5' GA GGA GAA GCC CGG TCT ACCATGAATCAC C 3'

2.4.2 PCR reactions

For amplification of the candidate sequences, DNA amplification was performed in a reaction volume of 50 μ l consisting of 2 μ l genomic DNA (from 114 ng/ μ l) from *C. difficile* strain 630 as template, 2.5 μ l of each primer (0.5 μ M in final concentration), 1 μ l deoxyribonucleoside triphosphates (200 μ M of each, final concentration), 0.5 μ l of Phusion polymerase (New England Biolabs) and 10 μ l of 5x Phusion buffer. Reactions were carried out for 35 cycles consisting of denaturation at 94°C (30 sec), annealing at 55°C (30 sec) and extension at 72°C (1 min) in a ThermoHybrid PX2 thermocycler. Table 2.2 shows a summary of the PCR reactions to show volumes and final concentrations. Amplified products were analyzed for size and yield *via* gel electrophoresis and those with the predicted size (see Table 2.3) were purified by QIA quick PCR purification kit (QIAGEN) for processing and cloning.

Table 2. 2. PCR reactions

COMPONENT	50 μ L REACTION	20 μ L REACTION	FINAL CONCENTRATION
H ₂ O	add to 50 μ l	add to 20 μ l	
5x Phusion HF or GC buffer	10 μ l	4 μ l	1x
10 mM dNTPs	1 μ l	0.4 μ l	200 μ M each
Sense primer	2.5 μ l	1 μ l	0.5 μ M
Antisense primer	2.5 μ l	1 μ l	0.5 μ M
Template DNA	2 μ l	0.8 μ l	
Phusion DNA polymerase	0.5 μ l	0.2 μ l	0.02 U/ μ l

2.4.3 Agarose gel electrophoresis

One percent agarose (Invitrogen) was prepared by adding 1 g of agarose powder into 100 ml of TEA (Tris-Acetate-EDTA; Amresco) buffer. The mixture was heated in a microwave until dissolved and cooled to 50°C. SYBR Safe (3µl; Invitrogen) was added to the cooled solution, it was mixed and poured into a gel casting tray prepared in advance with an appropriate comb for loading samples. After 15 minutes, the comb was removed and the tray was placed in the electrophoresis tank. TAE buffer was added. After mixing loading buffer (10X BlueJuice Gel Loading Buffer, Invitrogen) with the samples and a DNA ladder (100 kb or 1kb DNA ladder, Promega), samples were loaded into the gel. Electrophoresis was run at 100 V for 1 hour. DNA bands were visualised and recorded using a UV Pro Gold gel documentation system.

2.5 Cloning and expression of target proteins

2.5.1 Overview:

The pET-32 EK/LIC vector (Novagen) was used as the cloning and expression system. In ligation independent cloning system (LIC), inserts can be cloned directly into the plasmid without digestion or ligation. The plasmid itself was supplied with two overhanging termini which were complementary to overhanging sequences that were created on the PCR product. These sequences were included in primers used for amplification. To generate overhangs on the PCR products, they were firstly treated with T4 DNA polymerase. The 3' to 5' exonuclease activity of the enzyme created overhangs but in the presence dATP, the reaction ceased at a point determined by the primer sequence. The PCR products were then ready to anneal with complementary sequences carried on the vector (Figure 2.1).

The annealed DNA duplex was transformed into NovaBlue competent cells. These cells are not suitable for protein expression but can be used for DNA analysis to confirm the accuracy of the construct. Once accomplished (*eg* by colony PCR and sequencing), the plasmid can be transformed into *E. coli* strains such as BL21 (DE3) and BL21 (DE3)pLysS that are suited to expression of the recombinant protein. In pET-32 EK/LIC, recombinant proteins can be expressed with Trx and S tags for immunochemical detection, His tags for purification, and the inclusion of thrombin and enterokinase sites enable removal of these fused sequences if that is deemed important (Figure 2.2).

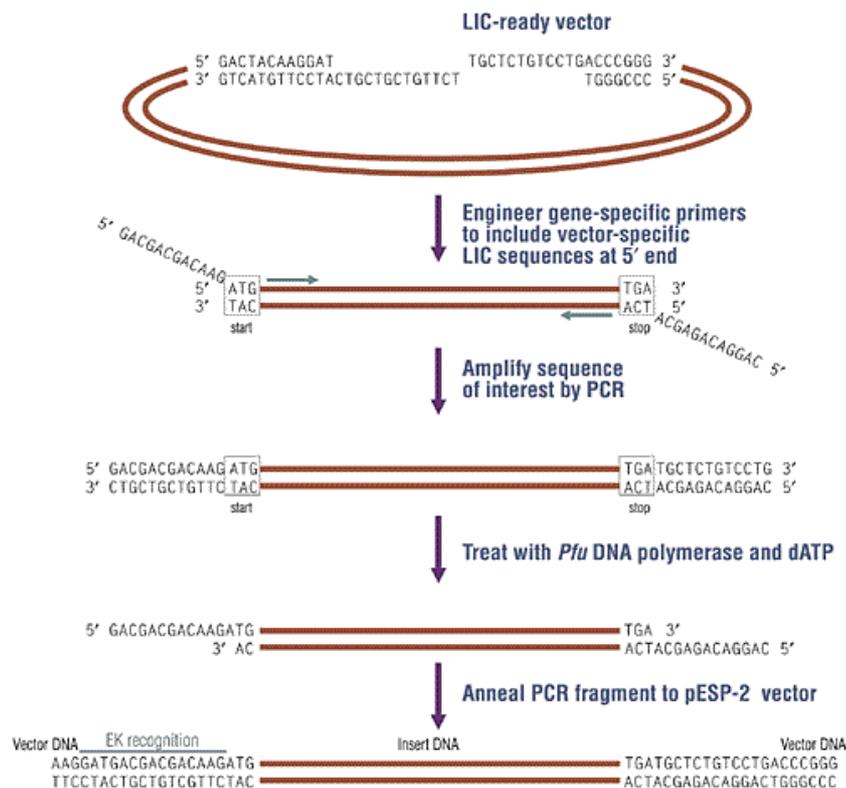
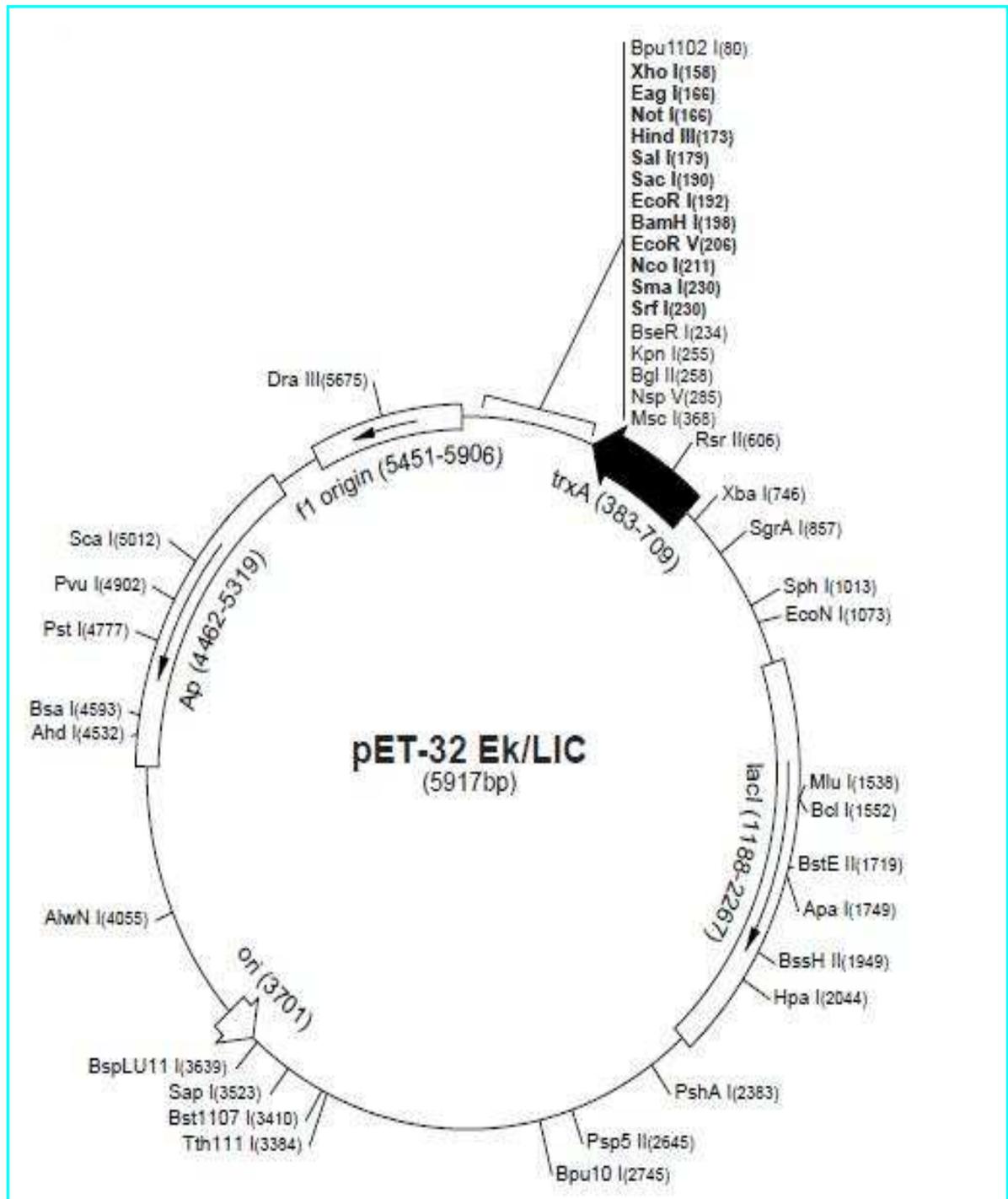


Figure 2. 1. LIC strategy showing the structure of the vector, generation and processing of a PCR product, and annealing of vector and insert. Adapted from data from Novagen.



2.5.2 Cloning protocol

2.5.2.1 T4 DNA polymerase treatment of purified PCR products

Purified PCR products were treated by T4 DNA polymerase of a quality certified for LIC in order to create the overhangs required for cloning. To do this, DNA concentrations were measured (Table 2.3) and 0.2 pmol of purified PCR product was added with other reactants to create a reaction of 20 μ l (Table 2.4). Reactions were started by adding T4 DNA polymerase and slowly pipetting the reaction mixtures for 30 minutes at 22°C. Reactions were then terminated by incubation at 75°C for 20 minutes. Prepared inserts were stored at -20°C in preparation for annealing to the pET-32 EK/LIC vector. As controls, a *lacZ* PCR product of 1815 bp was provided in the kit (100ng / μ l) as a positive control, and negative control reactions were assembled that lacked a PCR product (nuclease free water).

Table 2. 3. Purified PCR products and their concentrations

SAMPLE	BASE PAIR	PURIFIED PCR PRODUCT(NG/ μ L)	0.2 PMOL PURIFIED PCR IN NANOGRAM	VOLUME PURIFIED PRODUCT(μ G)	FROM PCR
<i>cwp84</i>	2265	62.4	294.45	4.8	
<i>cspA</i>	192	85.05	24.96	0.3	
5' <i>cwp66</i>	742	100.83	96.46	0.96	
<i>fbpA</i>	1689	82.6	219.57	2.66	
<i>groEL</i>	1461	74.46	194.61	2.62	
<i>fliD</i>	1497	61.14	194.61	3.14	
<i>acd</i>	1712	61.98	222.56	3.6	
<i>fliC</i>	780	88.6	101	1.5	
3' <i>cwp66</i>	898	69.38	116.74	1.65	
<i>sortaseB</i>	584	70.73	75.92	1.07	
<i>cwp66</i>	1700	43.84	221	5.04	

Table 2. 4. Components of T4 DNA polymerase reaction

VOLUME (μ L)	COMPONENT
x	0.2 pmol purified PCR product in TE buffer (10 mM Tris HCl, 0.1 mM EDTA, pH 8.0)
2	10X T4 DNA Polymerase buffer
2	25 mM dATP
1	100 mM DTT
y	Nuclease-free water to 20 μ l final volume
0.4	2.5 U/ μ l T4 DNA Polymerase LIC-certified

2.5.2.2 Annealing of treatment insert and pET-32 EK/LIC

The vector pET-32 EK/LIC was supplied by the manufacturer at 50 ng/ μ l ready for the annealing reaction. Fifty ng of vector DNA was added to 2 μ l of the T4 DNA polymerase reaction and incubated at 22° C for 5 minutes. Then, 1 μ l of 25 mM EDTA was added to the mixture for a further 20 minutes at 22° C. Samples were then used directly for transformation without further treatment.

2.5.3 Transformation

NovaBlue competent cells (Novagen) were used for transformation. A pre-prepared 50 μ l aliquot of NovaBlue competent cells was placed on ice and left to thaw for 5 minutes. Thawed tubes were flicked gently to resuspend the cells and then 1 μ l of annealed DNA duplex was added to each tube and stirred lightly. To check transformation efficiency an extra transformation was prepared containing closed circular plasmid DNA. Tubes were incubated on ice for another 5 minutes and heated for 30 seconds in a 40° C water bath without shaking. They were all put back on ice for an extra 2 minutes. Before plating on selective medium, 250 μ l of SOC medium was added to each tube and shaken (250 rpm) at 37° C for 60 minutes. Aliquots of 25 μ l were taken from each sample and plated to selective agar (2xYT ampicillin. 2xYT: tryptone 16 g/l, yeast extract 10 g/l, NaCl 5g/l, agar 15 g/l, mixed and autoclaved. Once cooled, ampicillin was added to 50 μ g/ml final concentration) for growth overnight at 37° C. Next day, all plates were checked for colonies.

2.5.4 Colony PCR and DNA sequencing

2.5.4.1 Colony PCR

Ten colonies from each transformation plate were identified for analysis. All 10 colonies were picked and inoculated to fresh 2x YT ampicillin for growth overnight at 37° C and then stored at 4° C for future use. Frozen stocks were also prepared for

each transformant by pelleting cells from 1 ml of overnight culture (liquid 2xYT ampicillin) and resuspending in 100µl of fresh liquid medium containing glycerol at 15% (v/v) final concentration. Colonies chosen from analysis were also picked and added to separate 50µl aliquots of sterile water for colony PCR. The tubes were placed in boiling water for 5 minutes to disrupt the cells and inactivate DNases. Then all tubes were centrifuged at 12000 x g for 1 min and 2 µl samples from the supernatants were used in PCR. PCR reactions were conducted as described earlier with primer sets required for each target sequence except that Taq DNA polymerase (GoTaq Flexi DNA polymerase, Promega) was used. Table 2.5 shows components of the colony PCR reactions. PCR products were analysed on 1% agarose gels as explained earlier to assess yield and the size of the products.

Table 2. 5. Components of PCR reactions for colony PCR

COMPONENT	25µL REACTION	FINAL CONCENTRATION
H ₂ O	add to 25 µl	
5X GoTaq Flexi Buffer	5 µl	1x
MgCl ₂ 25mM	2 µl	1mM
dNTPs 10 mM each	0.5 µl	0.2mM each
Sense primer	1.25 µl	0.5 µM
Antisense primer	1.25 µl	0.5 µM
Template	2 µl	
GoTaq DNA polymerase	0.25 µl	0.05 U/ µl

2.5.4.2 DNA sequencing

Based on the results of colony PCR, transformants from each cloning reaction were chosen and added to 2 ml liquid 2x YT-ampicillin medium and incubated overnight in a shaker (250 rpm) at 37°C. Plasmid DNA was isolated from the cells in each culture using a QIA prep spin miniprep kit (QIAGEN). Plasmid DNA was then sent for sequencing with T7 terminator primer (Figure 2.2).

2.6 Preparation of S-layer extracts using low pH glycine buffer

The method was provided by Professor N. Fairweather (Imperial College, London) and based on experience in his laboratory, it is estimated that a typical 50 ml culture of *C. difficile* will provide 0.5 – 1 mg of total surface protein including the high and low molecular weight derivatives of SlpA, the major protein of the S-layer. *C. difficile* overnight cultures were prepared in BHI broth under anaerobic conditions and cultured at 37°C for no longer than 18 hours. The *C. difficile* cells were pelleted by centrifugation at 5000 x g for 10 minutes at room temperature. Supernatant was removed and discarded into 2% Virkon to kill bacterial cells. The pellet was gently resuspended in 1/10 the initial culture volume using phosphate buffered saline (PBS) and centrifuged again. Washing in PBS was carried out three times to remove proteins and other material loosely associated with the bacterial surface. After the final wash, the bacterial pellet was drained onto paper and resuspended in 1/100 volume of low pH glycine (0.2 M glycine-HCl, pH 2.2) and incubated with gentle agitation at room temperature for 20 min.

The bacterial suspension was then centrifuged at 12000 x g in a bench top centrifuge at 4°C for 10 min. The supernatant containing proteins extracted from the surface was removed to a clean tube and its pH was brought to neutrality by

adding 2M Tris-base step-wise until the pH was 7.5 as estimated with a broad range pH indicator strip. The protein concentration of the extract was assayed using a bicinchoninic acid kit (Sigma). The composition of the extract was analysed by SDS-PAGE.

2.7 Flagella preparation

C. difficile was plated on four plates of Brazier's agar and kept for 48 hours at 37°C in an anaerobic cabinet. Bacteria were collected by scraping from the plates into 1 ml PBS. The suspension was agitated vigorously using a vortex mixer. The bacterial suspension was centrifuged at 5000 x g for 5 minutes to pellet whole cells and the supernatant containing extracted flagella was stored at -20°C pending future analysis.

2.8 Protein expression and purification of target proteins

In overview, to express recombinant proteins from the pET-32 EK/LIC vector, plasmid DNA was purified from NovaBlue cells and transformed into competent *E. coli* BL21 expression strains, specifically BL21 (DE3) or BL21 (DE3)pLysS. After selection on ampicillin medium and confirmation, the strains were used to express the recombinant proteins. Bacteria were grown in liquid culture, induced by addition of IPTG and then harvested by centrifugation. The cells were ruptured by using a sonicator to prepare a lysate for each sample. Cloning into the expression vector provided each recombinant protein with a histidine tag which was used for purification by Ni-affinity chromatography. Purified proteins were analysed by SDS-PAGE and Western blotting to check yields and molecular weight. To confirm the identity of the recovered protein(s) a single band was cut from SDS-PAGE gels and sent for analysis by mass spectrometry.

2.8.1 Transformation of plasmids into expression strains

Once the cloning of each target sequence into the expression vector had been confirmed, a single colony of NovaBlue cells carrying the plasmid was picked to

2ml of 2xYT with ampicillin and the culture was shaken at 250 rpm overnight at 37°C. The overnight cultures were centrifuged and the pellet was used to prepare plasmid DNA using QIA prep spin miniprep kits (QIAGEN). To transfer the purified plasmid into the expression strains, 20 µl aliquots of competent *E. coli* BL21(DE3) or BL21(DE3)pLysS cells (Novagen) were thawed on ice for 10 minutes. One µl of each plasmid was added to a sample of the thawed, competent cells and the mixture was placed on ice for 5 minutes after mixing very carefully. The cells were shocked by heating to 42°C for 30 seconds in water bath. They were returned to ice for 2 minutes.

To each transformation, 80 µl of SOC medium, pre-warmed to room temperature, was added and then the mixtures were shaken at 250 rpm at 37°C for 1 hour. Aliquots of 25 µl from each tube were plated on 2xYT-ampicillin and incubated at 37°C overnight. The next day, plates were checked for the appearance of colonies. A single colony from each plate was added into 200 µl of 2xYT-ampicillin and the culture was shaken at 250 rpm overnight at 37°C. For convenience and ease of storage, these cultures were prepared in 96 plates rather than individual tubes and from each transformation, 10 separate colonies were picked. After overnight culture, the 96 well plate was centrifuged at 1800 x g and the bacterial pellet in each well was resuspended in 2xYT-ampicillin containing glycerol at a final concentration of 15% (v/v). Plates were stored at -70°C to create a strain collection.

2.8.2 Protein expression

Ten µl of each bacterial stock was added to 200 ml of 2xYT-ampicillin (50µg/ml) and shaken at 180 rpm at 37°C until the culture reached an absorbance of 0.8 when measured at 600 nm. The cultures were induced by addition of IPTG (Sigma) to a final concentration of 1mM and shaken overnight at the lower temperature of 30 °C. Cells were collected from each overnight culture by spinning at 3000 x g for 10 minutes. The pellet was resuspended in 10 ml PBS and sonication was applied

to rupture the cells. To do this, the tube of suspended cells was placed on a plastic beaker containing ice making sure that the tube was completely immersed in ice to ensure thorough cooling. The sonicator probe was dipped into the cell suspension and activated for 10 seconds followed by a 30 seconds period of cooling. This was repeated 10 times. The probe was cleaned with 70% ethanol before moving to the next sample. After sonication, ruptured cells were centrifuged at 3000 x g for 10 minutes to remove the bacterial debris. The supernatants were frozen at -20°C pending further processing.

2.8.3 Protein purification

Ni affinity chromatography was used to purify the recombinant protein from each lysate. Pre-packed 5ml columns (Hi-Trap, GE Healthcare) were used. Before running a column, different solutions were prepared: binding buffer (20 mM phosphate, 0.5M NaCl, pH 7.4); binding buffer at 10x for adjustment of the pH and salt concentrations of lysates before purification; elution buffers (binding buffer containing imidazole at concentrations ranging from 40 mM to 0.5 M) to wash proteins from the columns and to elute bound proteins; nickel solution (0.1M NiSO₄); 50 mM EDTA to prepare columns for regeneration; 20% (v/v) ethanol for preserving the columns during long term storage at 4-8 °C.

To prepare the column, it was washed with 50 ml deionised water using a peristaltic pump (1-5 ml/min flow rate), and charged with 2 ml of NiSO₄ solution. The column was washed again with distilled water to remove any unbound Ni ions and equilibrated with 50 ml binding buffer. Bacterial lysate was mixed with 1/10 volume of 10x binding buffer and checked that the pH was adjusted to 7.4. The lysate was passed through the column, washed with 50 ml binding buffer, and then 25 ml of 40mM imidazole in binding buffer was pumped through to remove unbound or weakly-associated protein. Protein attached to the column was eluted with stepped imidazole beginning with 100 mM and finishing at 500 mM. Fractions were collected and analysed by SDS-PAGE and Western blotting. Single candidate

bands corresponding to the recombinant protein under analysis were cut from each gel and sent for analysis by mass spectrometry.

The fractions that contained recombinant protein of the predicted molecular weight were pooled and dialysed against 20 mM phosphate buffer, pH 7.4. After each round of purification, the column was washed with 50 ml distilled water and then treated with 5 ml 50 mM EDTA to remove all Ni ions. Newly regenerated columns were recharged with Ni, washed with distilled water and equilibrated with binding buffer, as described earlier.

2.9 Protein electrophoresis

Electrophoresis was carried out by two methods: SDS-PAGE and native PAGE. SDS-PAGE was carried out according to the Laemmli method. In contrast, native PAGE was run to analyse proteins in their native conformation using a discontinuous buffer system. In most cases, gels were hand cast and run in the mini-PROTEAN 3 system (BioRad). In some experiments, pre-cast gels of different acrylamide concentrations (Invitrogen) were used.

2.9.1 SDS-PAGE analysis

Hand cast gels were prepared as separating (typically 10% acrylamide; Table 2.6) and stacking gels (5% acrylamide). Twenty μl of sample buffer containing β -mercapthoethanol was added to 20 μl of the sample and heated to 95°C for 5 minutes. Gels were loaded with 10-15 μl of prepared sample and a sample of protein markers. After loading, gels were run at 120 volt until the gel front approached the bottom edge of the gel. Electrophoresis was terminated and the gel was stained with Coomassie brilliant blue solution for 1 hour. The stained gel was transferred into destain solution and shaken gently overnight to remove excess stain and clear the background. All gels were scanned and the data was stored electronically.

2.9.1.1 Stock solutions and buffers for SDS-PAGE

The following solutions were prepared: 30% acrylamide mixture containing 87.6 g of acrylamide and 2.4 g of bis-acrylamide dissolved and made up to 300 ml with deionised water (alternatively, a pre-prepared commercial acrylamide / bis solution, BioRad); 10% (w/v) SDS; 1.5M Tris-HCl, pH 8.8; ; 0.5M Tris-HCl, pH 6.8; sample buffer (3.55 ml deionised water, 1.25 ml 0.5M Tris-HCl pH 6.8, 2.5 ml glycerol, 2 ml 10% SDS and 0.2 ml of 0.5% (w/v) bromophenol blue to 9.5 ml final volume. Fifty μ l of β -mercaptoethanol was added to 950 μ l of sample buffer before use); 10x electrode (running) buffer pH 8.3 containing 30.3 g Tris base, 144 g glycine and 10 g SDS, brought to 1000 ml with deionised water; 10% (w/v) ammonium persulphate; staining solution (0.25% (w/v) Coomassie brilliant blue in a solution of 50% (v/v) methanol, 10% (v/v) acetic acid in water) and destain (15% (v/v) methanol, 10% (v/v) acetic acid in water). Table 2.6 shows the mixtures prepared for separating and stacking gels.

Table 2. 6. Solutions of SDS-PAGE separating and stacking gels

SOLUTIONS	10% SEPARATING GEL	12% SEPARATING GEL	5% STACKING GEL
H ₂ O	2ml	1.7ml	1.4ml
30% acryl amide solution	1.7ml	2 ml	0.33ml
1.5M Tris pH 8.8	1.3ml	1.3ml	----
1M Tris pH 6.8	----	----	0.25ml
10% SDS	0.05	0.05ml	0.02ml
10% APS	0.05ml	0.05ml	0.02ml
TEMED	0.002ml	0.002ml	0.002ml
Total volume	5 ml	5ml	2ml

2.9.2 Native PAGE analysis

Solutions for native PAGE were practically the same as prepared for SDS-PAGE except SDS was omitted. Resolving gels contained 8% or 10% acrylamide, topped with a 4% stacking gel (Table 2.7). Twenty μ l of sample was mixed with a native sample buffer (20 μ l) and heated to 95°C for 5 minutes. Prepared gels were loaded with 10-15 μ l of samples and a size markers and run at 100 volt until the blue dye front approached bottom of the gel. Staining, destaining and documentation were carried out as described earlier.

Table 2.7. Solutions of native PAGE Separating and stacking gels

SOLUTIONS	8% SEPARATING GEL	10% SEPARATING GEL	4% STACKING GEL
H2O	4.8ml	4.2ml	3.1ml
30% acryl amide solution	2.7ml	3.3ml	0.65ml
1.5M Tris pH 8.8	2.5ml	2.5ml	----
1M Tris pH 6.8	----	----	1.25ml
10% APS	0.05ml	0.05ml	0.025ml
TEMED	0.005ml	0.005ml	0.005ml
Total volume	10.055ml	10.055ml	5.03ml

2.10 Western blot analysis

For immunoblotting, SDS-PAGE gels were run and proteins were transferred to a nitrocellulose membrane. To locate each target, specific primary antibody (*eg* against an attached epitope tag) was used to probe the blot before addition of a second antibody HRP conjugate and addition of a chromogenic substrate. Each membrane was scanned for documentation.

2.10.1 Western blot

In preparation for transfer of proteins from the SDS-PAGE gel to nitrocellulose, four pieces of Whatman paper, nitrocellulose membrane, 2 permeable pads and a plastic cassette were assembled. All were soaked in transfer buffer (7.2 g Tris base, 33.4 g glycine and 600 ml methanol in 3 l total volume) for 5 minutes. The transfer sandwich was constructed by putting a pad and two pieces of Whatman paper into place, applying the SDS-PAGE gel, and then carefully smoothing nitrocellulose membrane into place over the gel, avoiding bubbles. This was topped with two pieces of Whatman paper and another porous pad. So far as practical, all components were kept wet with transfer buffer during the assembly process. The transfer sandwich was packed inside a blotting cassette and put inside the electroblotting tank (BioRad).

The tank was placed over a magnetic stirrer in a cold room and filled with cold transfer buffer. Transfer was conducted at 90 volt for one hour to move proteins from the gel to the membrane. After carefully unpacking, the membrane was washed with PBS and was stained with Ponceau red (Sigma) to check that transfer had taken place. The membrane was blocked overnight by incubating with 2% (w/v) skimmed milk in PBS (MPBS) at 4 °C with gentle shaking. The blocked membrane was washed 3 times with PBS and probed for one hour at room temperature with the first antibody, diluted in 2% MPBS. Anti S-tag antibody (detection of recombinant proteins) and anti-c-myc antibody (detection of recombinant scFv antibodies) were diluted 1/5000.

The membrane was washed 3 times with PBS-0.1% (v/v) Tween 20 and the second antibody, an HRP conjugated antibody, was added in 2% MPBS for another hour at room temperature. Anti mouse-HRP was used if the first antibody was anti S-tag; anti rabbit-HRP was used if the first antibody was anti-c-myc. In each case, the antibodies were diluted 1/5000. After final 3-fold washes with PBS-0.1% Tween 20, TMB substrate solution (Promega) was added and colour was allowed to develop. The reaction was stopped by washing with water.

2.11 Tomlinson library

The human single fold scFv libraries Tomlinson I and J were constructed at the MRC Centre for Protein Engineering and obtained for the study under a Material Transfer Agreement. These libraries were used for selecting scFv antibodies able to bind to the recombinant clostridial target proteins *via* phage display. Each library (I or J) contained over 100 million different scFv fragments integrated into a phagemid vector containing an ampicillin resistance gene. VH and VL domains of each scFv were connect together by a flexible glycine-serine linker sequence and the antibody coding sequence was followed by a His-tag for purification of scFv protein, a c-myc tag for immunochemical detection and an amber codon at the junction of the scFv coding sequence and the gene for protein III of the M13 phage coat.

Phage from each library with the capacity to bind target proteins were isolated by incubation in plastic tubes previously coated with the protein target of interest (panning). Unbound phage were then washed out; phage attached to the selection surface were eluted and amplified by infecting into *E. coli* TG1 cells. Three rounds of selection were used to enrich for phage binders. Screening was carried out to confirm the specificity of binding through ELISA. Soluble scFv were expressed from clones of interest and purified using the His-tag.

The Tomlinson libraries I and J libraries are constructed on single human frameworks for the VH heavy chain (V3-23/DP-47 and JH4b) and the V κ light chain (O12/O2/DPK9 and J κ 1). The libraries are diversified at 18 defined positions, specifically heavy chain residues H50, H52, H52a, H53, H 55, H56, H58, H95, H96, H97, H98, and light chain residues L50, L53, L91, L92, L93, L94 and L96. In library I, 95% of clones carry full-length scFv inserts; this is rather lower for library J (88%).

2.11.1 *KM13 helper phage*

2.11.1.1 Production of helper phage

Two hundred μl aliquots of *E. coli* TG1 grown to an OD of 0.4 were infected with 10 μl of 100-fold serial dilutions taken from a stock of KM13 helper phage (estimated at 10^7 pfu/ml). The mix of bacteria and virus were incubated in a 37°C water bath for 30 minutes without shaking. Three ml of molten H-top agar held at 42°C was added to each dilution and the mixture was poured onto pre-warmed TYE plates without antibiotic (TYE: 15 g agar, 8 g NaCl, 10 g tryptone, 5 g yeast extract in 1 l). After setting, the plates were incubated at 37°C overnight. A small plaque was picked from a plate which showed separated plaques and added into 5 ml fresh *E. coli* TG1 grown to 0.4.

The infected mix was grown for 2 hours at 37°C with shaking. Five ml of the culture was added into 500 ml 2xTY (2xTY: 16 g tryptone, 10 g yeast extract, 5 g NaCl in 1 l) in a 2 litre flask and grown with shaking at 37°C for 1 hour. Then kanamycin was added to a final concentration of 50 $\mu\text{g}/\text{ml}$ to select for infected cells. The culture was incubated at a 30°C with shaking overnight. The overnight culture was centrifuged at 10,800 x g for 15 minutes. One hundred ml PEG/NaCl (20 % (w/v) polyethylene glycol 6000, 2.5 M NaCl) was added to 400 ml culture supernatant and left for 1 hour on ice to precipitate the virus.

The prepared suspension was centrifuged at 10,800 x g for 30 minutes. The pellet was resuspended in 8 ml PBS and 2 ml PEG/NaCl was added and mixed to reprecipitate virus. After holding for 20 minutes on ice, it was centrifuged at 3300 x g for 30 min. The pellet was spun for the second time to remove any remaining residues of PEG/NaCl. The viral pellet was resuspended in 5 ml PBS and centrifuged at 11,600 x g for 10 minutes in a micro centrifuge to remove any remaining bacterial debris. The prepared helper phage were stored in 4°C for short periods or in PBS with 15% (v/v) glycerol for -70°C long-term storage.

2.11.1.2 Titration of produced helper phage

To titrate the helper phage, 5 μ l of trypsin stock (10 mg/ml trypsin in 50 mM Tris-HCl, 1mM CaCl₂, pH 7.4) was added to 45 μ l helper phage and incubated at 37°C for 30 minutes. One μ l of trypsin treated phage was diluted in 1 ml PBS and five, 100-fold dilutions were made from this using PBS. Fifty μ l samples from the dilutions were added to 1 ml of fresh *E.coli* TG1 culture (OD 0.4). The mixtures were added to molten H-top agar and poured onto pre-warmed TYE plate as described earlier. Plates were incubated at 37°C overnight. The same procedure was run for phage untreated with trypsin. KM10 encodes a trypsin-sensitive pIII protein and hence the expectation was that the titre of trypsin-treated phage would be 10⁵-10⁸ lower than for untreated phage.

2.11.2 Growing the scFv libraries

Library stocks of bacteria obtained from the MRC were added into 200 ml of pre-warmed 2xTY containing 100 μ g/ml ampicillin and 1% glucose and shaken at 37° until the OD at 600 nm reached 0.4. Helper phage (2x10¹¹) were added into 50 ml of culture and the remainder was frozen in aliquots for use as a secondary bacterial stock of the library. Infected cultures were incubated static at 37°C in a water bath for 30 minutes to allow efficient infection. Cultures were centrifuged for 10 minutes at 3000 x g.

The pellet was resuspended in 100 ml of 2xTY containing 100 μ g/ml ampicillin, 50 μ g/ml kanamycin and 0.1% glucose and shaken overnight at 30°C to allow phage replication and incorporation of the scFv coding phagemids into virus that simultaneously displayed the scFv-pIII fusion protein at their surface. The overnight culture was centrifuged at 3300 x g for 30 minutes and then 20 ml of PEG/NaCl was added to 80 ml of supernatant to precipitate the virus. It was mixed well and incubated on ice for 1 hour, then centrifuged for 30 minutes at 3300 x g. The supernatant was removed. As described earlier, remaining PEG/NaCl was collected by a second spin and discarded. The final pellet was resuspended in 4 ml of PBS and centrifuged at 11,600 x g for 10 minutes to remove remaining bacterial

debris. Aliquots of prepared phage were stored at -70°C in PBS with 15% glycerol. These phage were used in selection against recombinant protein targets.

2.11.2.1 Titration of phage from cultured libraries

For titration of phage carrying the scFv coding phagemid, 1 μl of phage stock was diluted in 100 μl PBS, and a series of six, 100-fold dilutions were prepared. *E. coli* TG1 was grown to an absorbance of 0.4 and 900 μl aliquots were added into each dilution of virus. These were incubated at 37°C for 30 minutes. TYE plates containing 100 $\mu\text{g}/\text{ml}$ ampicillin and 1% glucose were set out and 10 μl of each dilution was sampled and spread for single colonies. Plates were incubated overnight at 37°C . On infection, phagemid from the virus is propagated in the bacteria as a plasmid and hence the conversion of bacteria to ampicillin resistance defines the number of virus in the stock. It was recommended a phage stock with 10^{12} - 10^{13} cfu/ml should be used for selection.

2.12 Selection of recombinant antibodies by phage display

2.12.1 Selection on immunotubes

In selection, scFv with high affinity against the targets (Table 2.3) were isolated from the display libraries by panning. Normally three rounds of selection is sufficient to isolate binders against each target.

An immunotube (Nunc) was coated overnight at 4°C with 4 ml of a purified clostridial target protein at a concentration of 100 $\mu\text{g}/\text{ml}$ in PBS. Next day, each tube was washed 3 times with PBS and filled to the brim with 2% MPBS as a blocking reagent. Tubes were left for 2 hours at room temperature to allow

blocking to take place. After 2 hours, the immunotubes were washed 3 times with PBS. Library phage (10^{12} to 10^{13}) were mixed in 4 ml MPBS and added into each tube. Tubes were firstly incubated at room temperature for 1 hour and then rotated using an under and over turntable rotator for another hour, again at room temperature. The supernatant of each tube was thrown away and each tube was washed extensively with PBS-0.1% Tween 20: in the first round of selection, 10 washes were used; for selection rounds two and three, 20 washes were used. The excess PBS was removed by inverting each tube on a tissue for a few seconds.

The bound phages were eluted by adding 500 μ l of trypsin-PBS (50 μ l of 10 mg/ml trypsin stock in 450 μ l PBS) and rotating at room temperature for 10 minutes. From this eluate, 250 μ l was added into 1.75 ml of fresh *E. coli* TG1 (0.4) for static incubation at 37°C for 30 minutes in a water bath. The remaining 250 μ l of eluate was stored at 4°C as reserve. Serial 10-fold dilutions were prepared from the infected *E. coli* and 10 μ l aliquots were spotted on TYE plates containing 100 μ g/ml ampicillin and 1% glucose and incubated at 37°C overnight. The number of colonies found the following day enabled estimates to be made of the number of phage recovered from the selection step.

The remaining infected bacteria were centrifuged at 11,600 x g for 5 minutes in a micro centrifuge. The cell pellet was resuspended in 50 μ l of 2xYT and plated onto TYE (100 μ g/ml ampicillin and 1% glucose) for overnight incubation at 37°C. From the resulting colonies, phage were prepared to take forward further rounds of selection. To do this, the bacterial colonies were resuspended from the plate by adding 5 ml of 2xYT 15 % glycerol and loosening the cells with a sterile spreader. One ml of the resuspended cells were stored at -70°C as a stock, 50 μ l of were added into 50 ml of 2xYT (100 μ g/ml ampicillin and 1% glucose) and grown at 37°C until the absorbance reached 0.4. Ten ml were recovered and 5×10^{10} KM10 helper phage were added for static incubation at 37°C in a water bath for 30 minutes. The mix was then centrifuged at 3000 x g for 10 minutes and the pellet was resuspended in 50 ml 2xYT (100 μ g/ml ampicillin, 50 μ g/ml kanamycin and 0.1%

glucose) and shaken overnight at 30°C. The overnight culture was centrifuged at 3300 x g for 15 minutes and 10 ml of PEG/NaCl was added to 40 ml of the culture supernatant. To precipitate the phage, the suspension was mixed and placed on ice for 1 hour. After spinning at 3300 x g for 30 minutes, the supernatant was discarded and the pellet was spun again to remove any remaining PEG/NaCl. The viral pellet was resuspended in 2 ml PBS and again spun at 11,600 g for 10 min to remove bacterial debris. One ml of this phage was titred to determine the phage yield and then used for the next round of selection.

2.12.2 Screening phage by ELISA

ELISA using antibodies directed against the phage coat was used to monitor the specificity of reaction between scFv-pIII fusions and the target proteins. Mixed populations of phage recovered after each round of selection were tested in “polyclonal phage ELISA”; by picking single bacterial clones and preparing phage, virus were also tested in “monoclonal phage ELISA”.

2.12.2.1 Polyclonal phage ELISA

ELISA plates (96 well; Nunc) were coated with recombinant target proteins using the same concentration and buffer as applied in the selection process. For coating, 100 µl of protein was added to each well of the ELISA plate which was incubated overnight at 4°C. Plates were then washed three times with PBS and blocked by addition of 200µl 2% MPBS per well for 2 hours at room temperature. After washing three times with PBS, 10 µl of PEG precipitated phages recovered from each round of selection were added to each well, diluted in 100µl 2% MPBS. Plates were incubated for 1 hour at room temperature to allow binding to take place. After this, plates were washed 3 times with PBS-0.1% Tween 20 and 100 µl of HRP-anti-M13 conjugate (GE Healthcare) was added to each well. The conjugate

was diluted 1/5000 in 2% MPBS prior to use. Plates were incubated for 1 hour and washed three times with PBS-0.1% Tween 20. As substrate for the colorometric reaction, 100µl of TMB (Promega) was added to each well and left at room temperature for 10 minutes to allow development of a blue colour reaction. The reaction was stopped by adding 50µl of 1M sulphuric acid and the intensity of the resulting yellow colour was measured at 650 nm in a plate reader.

2.12.2.2 Monoclonal Phage ELISA

Individual colonies were picked from titration plates after each round of selection and bacteria were added into 100µl 2xYT (100µg/ml ampicillin and 1% glucose) in 96 well plates. Plates were shaken at 250 rpm at 37°C overnight. Two µl was sampled from each well and transferred into wells of a fresh plate, each containing 200µl of 2xYT (100µg/ml ampicillin and 1% glucose). Remaining overnight culture was mixed with glycerol to a final concentration of 15% and frozen to archive the strains.

Newly inoculated plates were shaken at 250 rpm at 37°C for 2 hours and then 25µl 2xYT was added to each well, each aliquot supplemented with ampicillin and glucose as before, but also containing 10^9 KM10 helper phage. Plates were shaken for a further hour at 37°C then centrifuged at 1800 x g for 10 minutes. Each bacterial pellet was resuspended in 200µl 2xYT (100µg/ml ampicillin and 50µg/ml kanamycin). After shaking overnight at 250 rpm, the plates were centrifuged at 1800 x g for 10 minutes and 50µl of the supernatant from each well was sampled and used in ELISA reaction as explained above.

2.13 Expression and purification of soluble scFv antibodies

In the display vector, a TAG stop codon was present between the scFv sequence and the coding sequence for pIII. In *E. coli* TG1, this was suppressed and a glutamate residue was incorporated instead, thereby creating a scFv-pIII fusion for incorporation into phage. Transfection of phagemid DNA into *E. coli* HB2151, a non suppressor strain, resulted in translation stopping at the end of the scFv sequence, producing a soluble antibody protein. Soluble scFv can be detected in ELISA with anti c-myc reagents, Protein A-HRP or Protein L-HRP. The presence of the His-tag also allows convenient purification.

2.13.1 Expression of solubilised scFv

E. coli HB2151 was grown to an OD of 0.4 and then 200 µl aliquots were infected with phage eluted from the third round of selection. Infection was allowed to take place at 37°C for 30 minutes without shaking. Four 100-fold serial dilutions were prepared and 50µl from each was plated onto TYE (100µg/ml ampicillin and 1% glucose). At least 20 individual colonies were picked into 100µl 2xYT (100µg/ml ampicillin and 1% glucose) in the wells of a 96 well plate that was then shaken at 250 rpm overnight at 37°C. Two µl was sampled from each well of the overnight culture and transferred into the wells of a fresh plate, each well containing 200µl 2xYT (100µg/ml ampicillin and 0.1% glucose). Remaining overnight culture was mixed with glycerol and frozen as a stock. The new plate was shaken at 250 rpm at 37°C until the OD reached 0.9 (typically, about 3 hours). At this point, 25µl of 2xYT (100µg/ml ampicillin and 9mM IPTG) was added to each well and shaking was continued overnight at the lower temperature of 30°C. The following day, the plate was spun at 1800 x g for 10 minutes and 50µl of supernatant of each well was used ELISA.

2.13.2 Detection of soluble antibodies

To check for the successful production of soluble antibodies and their specificity, the wells of an ELISA plate were coated with 100µl of target protein, washed and blocked with 200µl of 3% BSA in PBS, following the protocol described earlier. After 2 hours, the plate was washed 3 times with PBS and 50µl of each test supernatant thought to contain scFv was added to test wells with 50µl of 3% BSA-PBS. Plates were incubated for 1 hour at room temperature, washed 3 times with PBS-0.1% Tween 20 and 100µl of rabbit anti-c-myc (Sigma) was added to each well. The antibody was diluted 1/5000 in 3% BSA-PBS and was allow to interact for 1 hour at room temperature.

After this, plates were washed 3 times with PBS-0.1% Tween 20 and 100µl of diluted goat anti-rabbit-HRP (Sigma) was added for another 1 hour at room temperature. The conjugate was diluted 1/5000 in 3% BSA-PBS prior to use. After washing the plate 3 times with PBS-0.1% Tween 20, 100 µl TMB substrate (Promega) was added, quenched and quantified as described earlier.

2.13.3 Purification of soluble antibodies

After selecting HB2151 clones which produced scFv with strong reaction against the target, large scale cultures were grown for the purification of solubilised recombinant antibodies. A 5ml sample of 2xYT (100µg/ml ampicillin and 1% glucose) was inoculated with a single colony of selected HB2151 for each scFv and was shaken overnight at 250 rpm at 37°C. The next day, 3ml of overnight culture was added to 300 ml of 2xYT (100µg/ml ampicillin) and shaken at 180 rpm at 37°C and monitored until the absorbance of the culture reached 0.9. scFv expression was then induced by adding IPTG to a final concentration of 1mM. Incubation then continued with shaking at 180 rpm for 18 hours at a temperature of 30°C. Bacterial cells were collected from the overnight cultures by centrifugation at 3000 x g at

4°C for 10 minutes. ScFv proteins were purified from the supernatant of each culture using Ni-affinity chromatography.

For purification, the following solutions were prepared: binding buffer (20 mM phosphate, 0.5M NaCl, pH 7.4); 10x binding buffer for adjusting the pH and salt concentrations prior to chromatography; imidazole at 20, 40, 200 mM in binding buffer; 0.1M NiSO₄; 50 mM EDTA for regenerating the column after purification. A 5ml pre-packed His-Trap column (GE Healthcare) was washed with 10 volumes of deionised water and charged by addition of 2ml of 0.1M NiSO₄ in preparation for binding the recombinant antibody *via* its His-tag. As described earlier, a peristaltic pump working at a flow rate of 5 ml/min or less was used to deliver buffer *etc* to the column. After charging with Ni, the column was washed with 50 ml deionised water and 50 ml of binding buffer.

One hundred ml of supernatant from induced cultures was adjusted for pH by adding 10x binding buffer and was then pumped onto the equilibrated column. The loaded column was washed with 25 ml of binding buffer followed by 25 ml of 40 mM imidazole in binding buffer to remove unbound and loosely associated proteins from the column. Initial experiments explored elution with imidazole ranging in concentration from 100 to 500 mM and from this, a standard protocol was developed in which scFv proteins were eluted with 200 mM of imidazole in binding buffer. Fractions were collected during elution and each was analysed by SDS-PAGE and also dot blotting. Evidence of a 30 kDa band and positive reactions in dot blot identified fractions to be pooled and dialysed against PBS in preparation for further use. After purification of each scFv, the column was washed with 50 ml distilled water and treated with 5 ml of 50 mM EDTA to remove Ni ions. The column was then regenerated by charging with a new aliquot of NiSO₄, washing with deionised water and equilibrating with binding buffer.

2.14 Immunochemical characterisation of recombinant antibodies

To assess the ability of scFvs to recognise the recombinant clostridial proteins, Western blot analysis was conducted with target proteins after separation by native and denaturing gel electrophoresis. These experiments were based on protocols for SDS-PAGE and native electrophoresis described earlier. The target proteins (*eg* SlpA extracted from *C. difficile* with glycine buffer) were separated by gel electrophoresis and transferred to nitrocellulose membrane. The membrane was probed with recombinant antibodies of the required specificity, anti-c-myc antibody to bind to the peptide tag on the scFv, anti-rabbit-HRP conjugate and TMB substrate as described earlier. For some experiments, gels were loaded with a range of proteins for probing (*eg* SlpA extracted from different *C. difficile* strains, flagellar preparations, and bacterial lysates). Experiments are detailed in the Results section of the thesis.

2.15 Immunostaining and electron microscopy of *C. difficile* cells using scFvs

Immunofluorescence microscopy was carried out with purified scFv to understand whether recombinant antibodies were able to recognise target proteins in their native state and in their normal location on the *C. difficile* cell. *C. difficile* cells were cultured overnight and applied to glass microscope slides by cytospin centrifugation. After blocking, the bacteria were probed with scFv against the target protein of interest, followed by Protein A-FITC to generate the fluorescent signal. Electron microscopy was also used in which target proteins were labelled with scFv and the location of the recombinant antibodies was determined with Protein A-gold, or alternatively anti-c-myc followed by anti-rabbit-gold conjugates.

2.15.1 Indirect immunofluorescence

A single colony of *C. difficile* was picked and inoculated to 10 ml aliquots of BHI medium, and grown overnight at 37°C under anaerobic conditions. The following day, the cultures were checked for growth to an approximate absorbance of 0.6. The cultures were centrifuged at 3000 x g for one minute and the pellets were washed three times with 1ml aliquots of PBS. Each 1 ml aliquot of cells suspension was derived from 10 ml overnight culture; each was sampled and 100 µl was used in the cytospin (Cytospin 3, Shandon; 500 rpm for 5 minutes) to attach the clostridial cells to microscope slides.

Alternatively, cell pellets were mixed with 1ml of 4% paraformaldehyde (PFA), resuspended and incubated for 15 minutes at room temperature to fix the bacteria. Treated cells were collected at 3000 x g for 1 minute and washed three times with PBS. To quench the remaining PFA, 1ml of 20 mM NH₄Cl₂ was added to each tube and incubated on the bench for 15 minutes. The cells were pelleted and again washed three times with PBS. The pellet from each tube was resuspended in 1 ml PBS and 100µl aliquots were used for each slide.

Bacteria on the slides were treated with a few drops of ethanol to fix the cells and they were then washed with PBS and blocked by adding 200µl of 2% BSA-PBS as blocking reagent. Blocking solution was incubated on the sample zone for 4 hours at room temperature inside a humidified environment. The slides were washed three times with PBS and 50µl of each purified scFv was applied, diluted to 1/50 in 0.1% BSA-PBS. Slides were incubated on room temperature for 1 hour.

After washing three times with PBS, the slides were probed with 50 µl of a 1/50 dilution of Protein A-FITC (Sigma) in 0.1% BSA-PBS. The conjugate was allowed to act for 1 hour at room temperature. The slides were washed three times with PBS and allowed to dry on the bench. Anti-fade agent (Dako; 10-20µl) was applied to the bacterial zone and covered with coverslip which was sealed into place with nail

polish. The slides were visualised with a fluorescence microscope (Axiovert 100, Zeiss).

2.15.2 Electron microscopy

Bacterial cells from *C. difficile* 630 were prepared as previously explained and the washed bacterial suspensions were used to prepare grids for electron microscopy. The bacterial pellets were first fixed with 4% paraformaldehyde-0.1% glutaraldehyde in PBS for 30 to 60 minutes on ice. They were then washed by gently adding and removing PBS twice for 5 minutes with each wash. The fixed bacteria were dehydrated through stepped ethanol solutions: initially, the pellet was incubated at room temperature in 50% ethanol for two exposures each of ten minutes; following this, three exposures of ten minutes each were in 70%, 90% and absolute ethanol, all at -40° C.

Embedding of specimens began with pre-treatment with different percentages of HM20 resin. The dehydrated bacterial pellet was added at -40° C to: 50%:50% (v/v) dried absolute ethanol and HM20 resin for a maximum one hour; 25%:75% (v/v) ethanol and resin overnight; overnight incubation in HM20 alone. After this treatment, the specimen was transferred to gelatine capsules at -40° C and exposed to UV (340 nm) for 24-48 hours to polymerise the resin. Sections of 60-70 nm were then cut onto nickel 300 mesh grids coated with formvar.

For immunolabelling, 40 µl of a 1/10 dilution of purified scFvs was added to clean layer of parafilm inside a chamber and grids were floated on the droplets for one hour. After washing three times with PBS, the grids were labelled with a conjugate of Protein A-5nm colloidal gold, diluted in 1% BSA in PBS. Alternatively, the grids were probed with 1/10 dilution of anti-c-myc in PBS, washed with PBS then probed with anti-rabbit-gold conjugate. Each incubation was for one hour and was followed with three washes with PBS.

All grids were treated with 2% aqueous uranyl acetate for 10 minutes and Reynolds lead citrate for 5 minutes. Grids were then checked in a LEO 912AB transmission electron microscope and images captured for further analysis.

2.16 Functional characterisation of anti-flagellar scFv antibodies

Many *C. difficile* strains are motile as a consequence of the multiple flagella on their surfaces. The recombinant scFv antibodies against FliD and FliC were checked in a motility inhibition test to assess if they were able to block this aspect of bacterial behaviour. Three *C. difficile* strains were used in this study: 630 and R20291 which are both motile; M120 which is non-motile and therefore served as a negative control. Further controls were the use of scFvs against SlpA, a protein known to be present at the bacterial surface, but a target that was unlikely to effect motility when scFvs were attached.

2.16.1 Counting of C.difficile from overnight culture

Ten ml of overnight cultures of *C. difficile* were prepared under anaerobic conditions at 37°C. The cultures were centrifuged at 3000 x g for 10 minutes and bacterial pellets were resuspended in 1 ml of PBS. Four, 10-fold serial dilutions were prepared and 50µl from each was plated on supplemented Brazier's agar, prepared and supplemented as described earlier. The plates were incubated at 37°C for 48 hours in an anaerobic cabinet and colonies were counted to establish the number of colony forming units in the original overnight liquid culture.

2.16.2 Motility inhibition test

To test the effect of recombinant antibodies on the movement of bacteria, aliquots were prepared, each containing 10⁵ *C. difficile* cells. These aliquots – roughly 10µl of overnight culture – were incubated with 150µl of supernatant from an induced culture of *E. coli* HB2151 and known to be expressing the anti-FliD or anti-FliC of interest. Bacteria were incubated with the crude scFv sample for 1 hour at 37°C in an anaerobic cabinet. Then, 10µl of the mixture was stabbed into 0.2%

BHI agar and kept at 37°C in the anaerobic cabinet and monitored for growth over 48 to 72 hours. Controls were set up with the test strains of bacteria but no scFv. The antibody LA6, a scFv against SlpA, was used in place of the anti-FliD, anti-FliC antibodies as a control that was not expected to block bacterial motility. Motility was evident as diffusion of bacteria out from the site of inoculation, moving through the low concentration agar.

3 Results

3.1 PCR for clostridial targets

As described earlier, a number of bacterial proteins were chosen as targets for the study based upon their known or likely surface location. Coding sequences were amplified by PCR from the genome of *Clostridium difficile* 630 for cloning, expression and purification as recombinant proteins. Carrying this out, two polymerases were tested, GoTaq from Promega, a conventional Taq DNA polymerase and Phusion DNA polymerase, a high fidelity enzyme from New England BioLabs. Amplifications were carried out in standard PCR reactions using gene-specific primers that carried sequences required for later cloning into the pET32 EK/LIC system (Table 3.1). Initially, products were analysed from the PCR reaction carried out with Taq DNA polymerase and genomic DNA (g-DNA) of *C. difficile* 630 as the template. Gels showed the amplification products were of the expected size for all samples with a good intensity and clear backgrounds.

In Figure 3.1, the second lane from the left hand side shows PCR products for *cwp84*. Based on primer annealing sites (Table 3.2) the expected size for the band is 2265 bp; a product of approximately this size can be seen when compared to the DNA ladder (lane 1). In Table 3.2, it is predicted that the expected size for the *cspA* product is 192bp. The migration of DNA in lane 3 in Figure 3.1 is in agreement with this prediction. For *cwp66*, two reactions were run as can be seen in lanes 4 and 10. These reactions recovered the 5' region of *cwp66* and its 3' coding sequence respectively. Considering the sizes calculated in Table 3.2 for 5' and 3' regions of *cwp66*, lane 4 in Figure 3.1 shows a single band close to 750 bp as expected and the amplification product seen in lane 10 is also of the expected size. The band in lane 5 in (Figure 3.1) is also of a size consistent with that predicted for *fbpA* (1689 bp, Table 3.2) as is

the amplification product for *groEL* (a band just under 1500 bp compared with the expected size of 1461 bp).

In Figure 3.1, lanes 7 and 8 show the PCR products for the *fliD* and *acd* reactions. The band in lane 7 for *fliD* is consistent with the predicted size (1497 bp). Table 3.2 predicts a product of 1650 bp band for *acd* and although a band can be seen in lane 8 (Figure 3.1) in the right position, an unexpected product of 1000 bp can also be seen. The PCR product for *fliC* was of the correct size (780 bp; lane 9 of Figure 3.1). In lane 11, the *sortaseB* amplification generated a band slightly over 500 bp in size. This is in agreement with the predicted size of 584 bp.

To reduce the chances of errors arising in the amplification products, PCRs were also carried out with Phusion DNA polymerase. This enzyme was supplied with two buffers, HF and GC, the manufacturer suggesting HF as the default buffer. Figure 3.2 shows the results of amplification with Phusion DNA polymerase using buffer HF; all products are of the predicted size except for *cwp84* and *fbpA*. Lane 2 shows the reaction for *cwp84* and, as detailed earlier, the reaction should yield a product of 2265 bp. Amplification with Phusion DNA polymerase band buffer HF generated a band of less than 500 bp. In lane 5, no band can be seen; the *fbpA* reaction was supposed to generate a product of 1689 bp (Table 3.2. When comparing predicted sizes of other PCR products with the actual data, Figure 3.2 lanes 3 (*cspA*), 4 (5' *cwp66*), 6 (*groEL*), 7 (*fliD*), 8 (*acd*), 9 (*fliC*), 10 (3' *cwp66*) and 11 (*sortase B*) shows that all products were in the right positions.

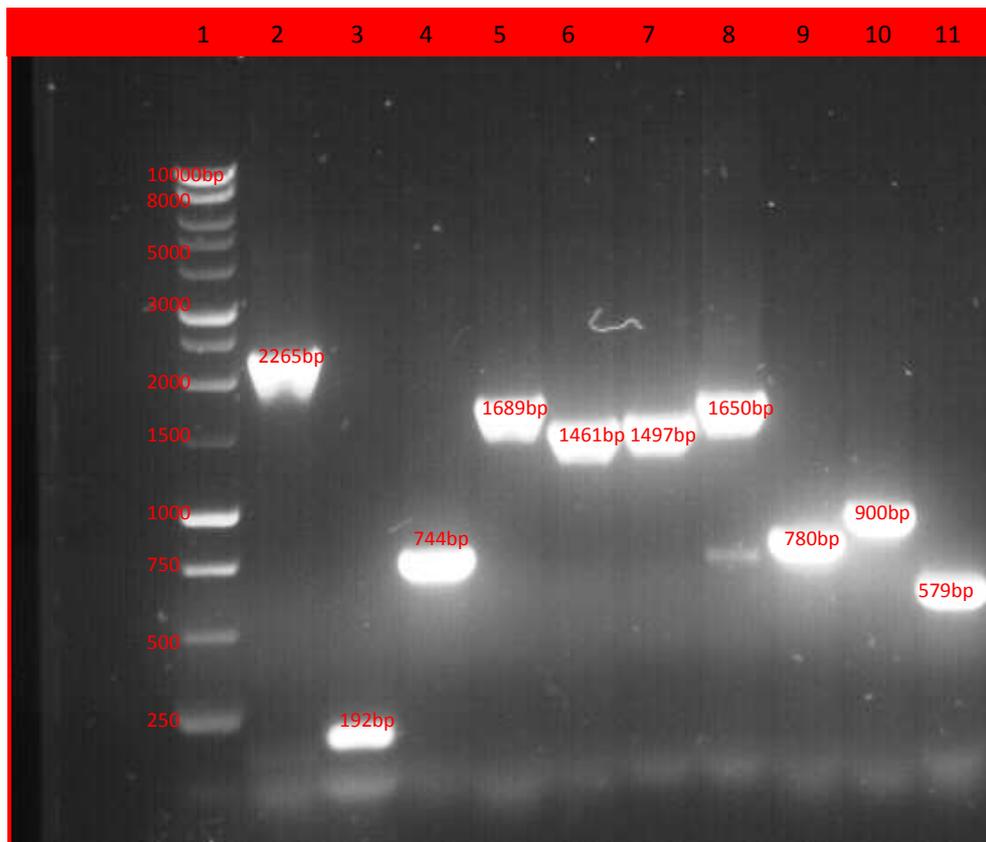


Figure 3.1. PCR products obtained with sense and antisense primers using Taq DNA polymerase (Promega).

From left: lane 1, 1kb DNA ladder(Promega); lane 2, *cwp84*; lane 3, *cspA*; lane4, 5' *cwp66*; lane 5, *fbpA*; lane 6, *groEL*; lane 7, *fliD*; lane 8, *acd*; lane 9, *fliC*; lane 10, 3' *cwp66*; lane 11, *sortaseB*.

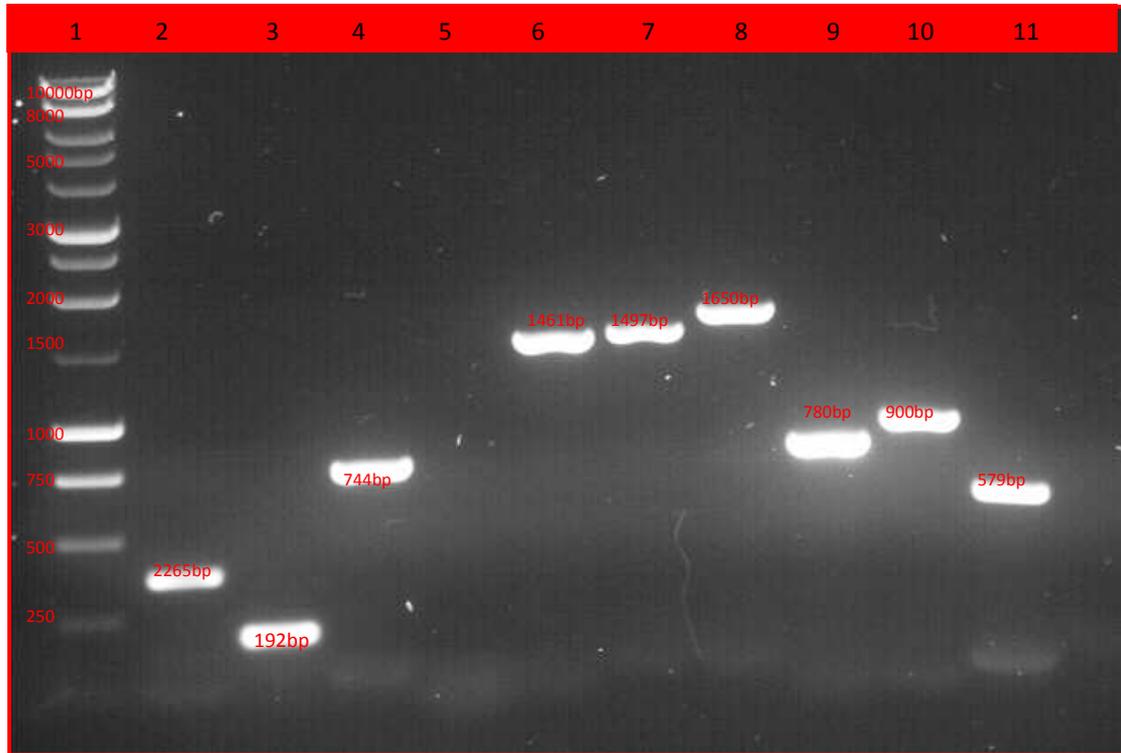


Figure 3.2. PCR products of targets with Phusion DNA polymerase (New England BioLabs).

From left: lane 1, 1kb DNA ladder (Promega); lane 2, *cwp84*; lane 3, *cspA*; lane 4, 5' *cwp66*; lane 5, *fbpA*; lane 6, *groEL*; lane 7, *fliD*; lane 8, *acd*; lane 9, *fliC*; lane 10, 3' *cwp66*; lane 11, *sortase B*.

Table 3. 1. Sequences of sense and antisense primers for amplification of *C. difficile* targets by PCR for ligation independent cloning.

TARGET NAME (CD)	PRIMERS SENSE (UP SIDE SEQUENCE) ANTISENSE(DOWN SIDE SEQUENCE)
<i>cwp84</i> (2787)	5' GAC GAC GAC AAG ATA GAT GGA GTA GAA ACT GCA GAG 3' 5' GA GGA GAA GCC CGG TTC ATT TCC ATT TCC ACC AAC 3'
<i>cspA</i> (0892)	5' GAC GAC GAC AAG ATG AAA AAC GGA ATA GTA AAA TGG 3' 5' GA GGA GAA GCC CGG TAC GTT TTC AGC TTG AGG TCC 3'
5' <i>cwp66</i> (2789)	5' GAC GAC GAC AAG ATA ACG GGT TCT GGA AGA TGG 3' 5' GA GGA GAA GCC CGG TTT AGC TGC TAA TAC ACC CAC 3'
<i>fbpA</i> (2592)	5' GAC GAC GAC AAG ATA CAT CAA CCT GAA GAT GAT GAG 3' 5' GA GGA GAA GCC CGG TTT AAC CTT AAG CTT GGC TAC 3'
<i>groEL</i> (0194)	5' GAC GAC GAC AAG ATTGGA GTA ACT ATA GCA AAA GAG 3' 5' GA GGA GAA GCC CGG TCC GCC ACC CAT TCC TGG 3'
<i>fliD</i> (0237)	5' GAC GAC GAC AAG ATT CCA GTA AGA GTT ACA GGC 3' 5' GA GGA GAA GCC CGG TTG TGA GAA ATA GTT CAT TTG 3'
<i>acd</i> (1304)	5' GAC GAC GAC AAG ATT GAA CCA ACT GCC GAA AGT AGC 3' 5' GA GGA GAA GCC CGG TTC CAT AAT TCC AGA AAT TCC 3'
<i>fliC</i> (0239)	5' GAC GAC GAC AAG ATG GAG AAG TTA TCT TCT GGG 3' 5' GA GGA GAA GCC CGG TAA AAC TCC TTG TGG TTG TTG 3'
3' <i>cwp66</i> (2789)	5' GAC GAC GAC AAG ATA GTT ACT CAA ATT GGT GGC 3' 5' GAC GAC GAC AAG ATA GTT ACT CAA ATT GGT GGC 3'
<i>sortaseB</i> (2718)	5' GAC GAC GAC AAG ATC AAT CAT GAT ACT AAA ATA TCC 3' 5' GA GGA GAA GCC CGG TCT ACCATGAATCAC C 3'
<i>cwp66</i> (2789)	5' GAC GAC GAC AAG ATA ACG GGT TCT GGA AGA TGG 3' 5' GAC GAC GAC AAG ATA GTT ACT CAA ATT GGT GGC 3'

In the primer sequences, bold text shows the complementary sequence of prepared LIC vector

Table 3. 2. Predicted size and molecular weight of the targets

Predicted sizes show the theoretical sizes of the PCR products in bp for each reaction (column 3), the anticipated size of proteins encoded by the PCR products (kDa) and data for proteins expressed from the pET32 EK/LIC expression vector(+tag kDa).

TARGET NAME	ACCESSION NUMBER	PREDICTED SIZE (bp)	PREDICTED MOLECULAR MASS OF PROTEIN (kDa)	PREDICTED MOLECULAR MASS + TAG (kDa)
<i>cwp84</i>	YP_001089300.1	2265	82.04	102.51
<i>cspA</i>	YP_001087366.1	192	7.06	27.53
<i>5' cwp66</i>	YP_001089302.1	744	26.6	47.07
<i>fbpA</i>	YP_001089109.1	1689	64.71	85.18
<i>groEL</i>	YP_001086664.1	1461	51.73	72.2
<i>fliD</i>	YP_001086705.1	1497	55.31	75.78
<i>acd</i>	YP_001087799.1	1650	59.75	80.22
<i>fliC</i>	YP_001086707.1	780	27.57	48.04
<i>3' cwp66</i>	YP_001087366.1	900	33.47	53.94
<i>sortaseB</i>	YP_001089230.1	579	22.91	43.38
<i>cwp66</i>	YP_001089302.1	1722	62.75	83.22

Further PCRs were run to assess the influence of buffers HF and GC on amplification of *cwp84*, *fbpA* and full-length *cwp66* sequences with Phusion DNA polymerase. In Figure 3.3, lanes 2 and 3 compare the outcome of PCR reactions for *cwp84* using Phusion DNA polymerase and buffers HF or GC respectively. Based on the data in Table 3.2, the amplified fragment should be 2265 bp, a result which is evident in lane 3 when GC buffer was used for the PCR reaction. Using HF buffer in the reaction resulted in a product too low in molecular weight (lane 3). The sample loaded to lane 6 was a *cwp84* amplification using Taq DNA polymerase. Lanes 4, 5 and 7 in Figure 3.3 show the reactions for *fbpA*. Phusion DNA polymerase was used for PCRs loaded to lanes 4 and 5, Taq DNA polymerase for lane 7. No product was obtained using HF Buffer (lane 4) but when GC buffer was chosen (lane 5), the *fbpA* reaction generated a product consistent with the predicted size (1689 bp). This result was also obtained with Taq (lane 7). The PCRs for *cwp66* were successful (1700 bp band) using Phusion DNA polymerase with GC buffer (lane 8) and Taq polymerase (lane 9). Again, no product was obtained from this reaction when HF buffer was used with Phusion DNA polymerase (Figure 3.3, lane 10).

Having optimised the reactions, further PCRs were carried out so that the products could be extracted from agarose gels for cloning. Eleven reactions were run with Phusion DNA polymerase to generate amplicons for the targets *cspA*, 5' *cwp66*, *groEL*, *fliD*, *acd*, *fliC*, 3' *cwp66*, putative *sortaseB* (HF buffer) and with GC buffer for *cwp84*, *fbpA* and the full length *cwp66* sequence. The reaction products were all loaded to agarose gels and the relevant bands were excised and purified using a gel extraction kit from QIAGEN. After extraction, the PCR products were all analysed again by agarose gel electrophoresis to re-check their size. Figure 3.4 shows that all 11 samples of extracted DNA contain products of a size consistent with that predicted in Table 3.2. The amount of DNA present in each extracted sample was estimated by nanodrop (Table 3.3) in preparation for cloning.

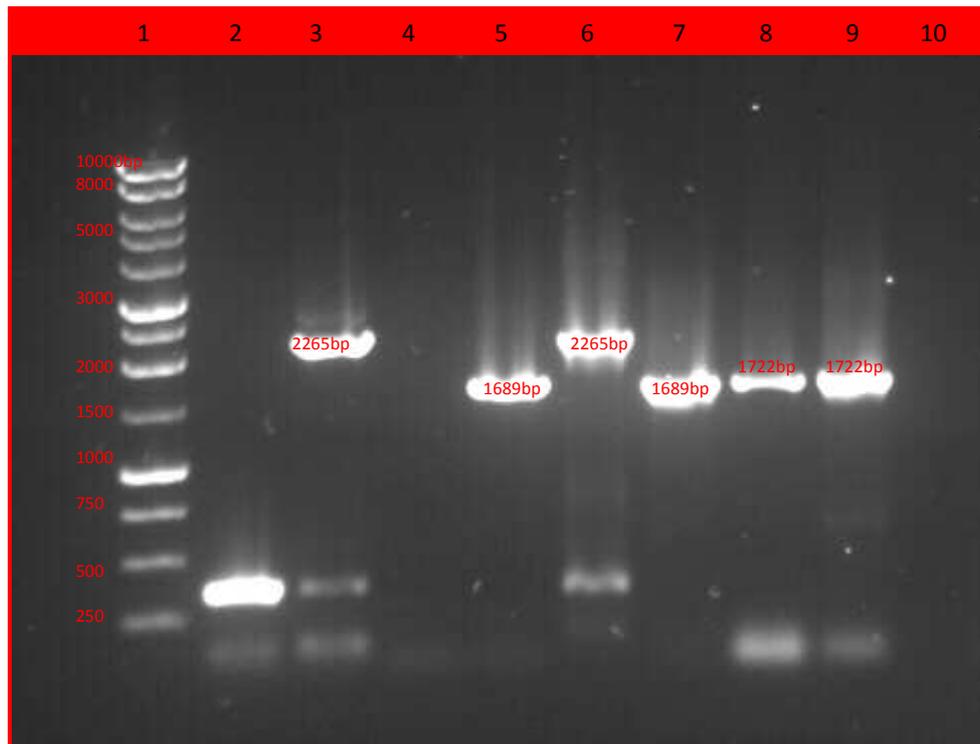


Figure 3.3. Comparison of PCR reactions using Phusion DNA polymerase with buffers HF and GC buffers, and Taq polymerase.

Lane 1, 1kb DNA ladder (Promega); lane 2, *cwp84* with HF buffer; lane 3, *cwp84* with GC buffer; lane 4, *fbpA* with HF buffer; lane 5, *fbpA* with GC buffer; lane 6, *cwp84* with Taq polymerase; lane 7, *fbpA* with Taq polymerase; lane 8, *cwp66* with GC buffer; lane 9, *cwp66* with Taq polymerase; lane 10, *cwp66* with HF buffer. Phusion DNA polymerase was used in all reactions in which HF or GC buffer was tested.

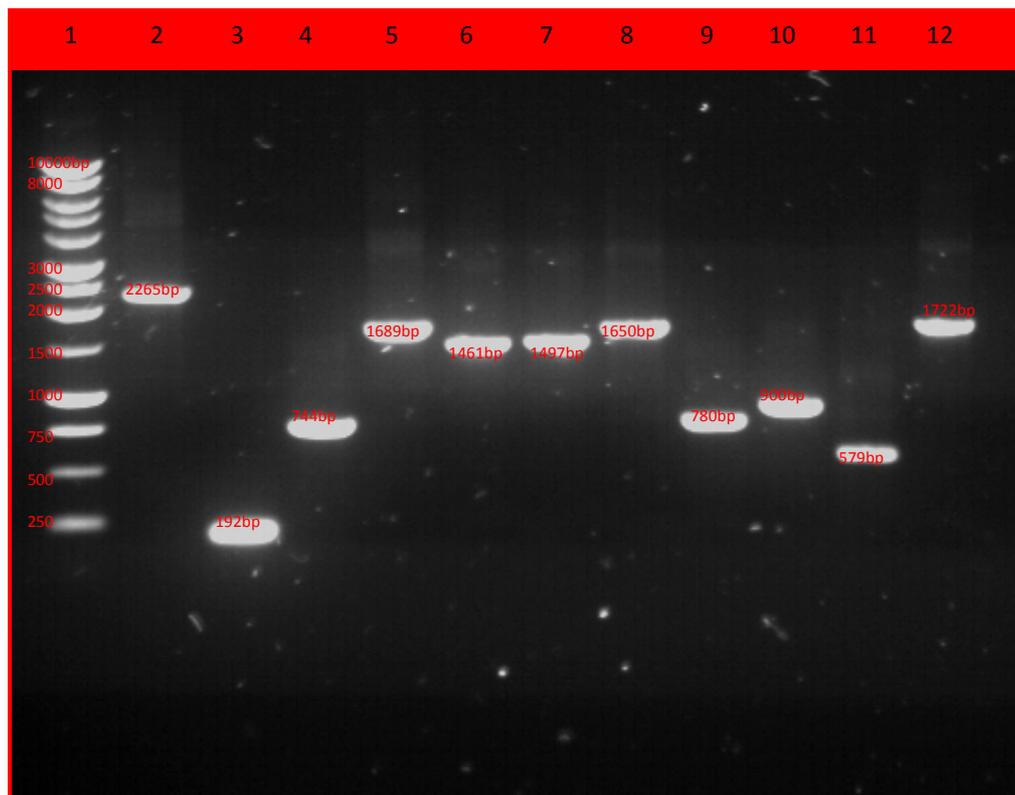


Figure 3.4. Agarose gel electrophoresis of extracted PCR products. Lane 1, 1kb DNA ladder; lane 2, *cwp84*; lane 3, *cspA*; lane 4, 5' *cwp66*; lane 5, *fbpA*; lane 6, *groEL*; lane 7, *fliD*; lane 8, *acd*; lane 9, *fliC*; lane 10, 3' *cwp66*; lane 11, *sortaseB*; lane 12, *cwp66*. All PCR reactions were carried out with Phusion DNA polymerase with buffer HF (lanes 3, 4, 6, 7, 8, 9, 10, 11) or GC buffer (for lanes 2, 5 and 12.)

Table 3. 3. DNA concentrations of extracted samples for use in cloning

TARGET	CWP84	CSPA	5'CWP66	FBPA	GROEL	FLID	ACD	FLIC	3'CWP66	SORTASE B	CWP66
ng/ μ l	62.40	85	100.8	82.6	74.46	61.14	61.98	88.60	69.38	70.73	43.84

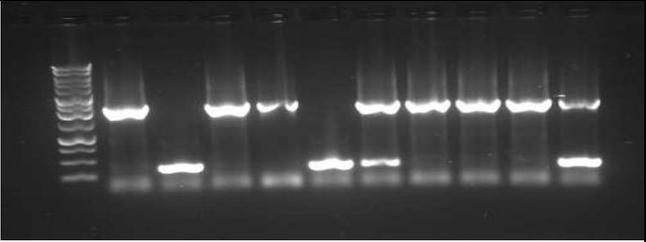
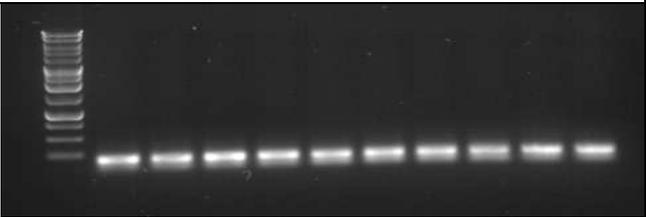
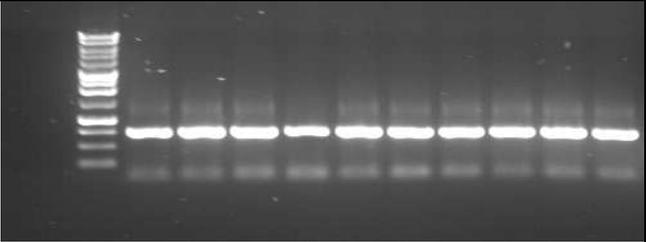
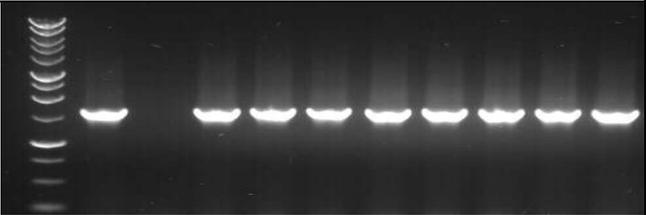
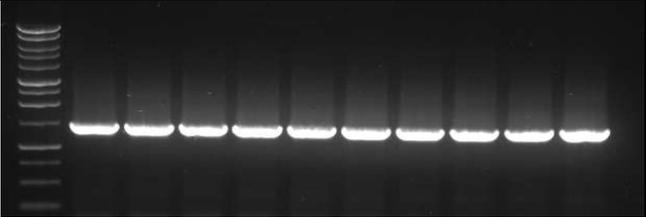
3.2 Cloning and expression of targets using LIC cloning system

Having optimised their amplification, PCR products were then used for cloning so that the target proteins could be expressed. This phase of the project entailed a series of different steps for each target. Each PCR product was first cloned into the pET-32 EK/LIC vector and transformed into a strain of *Escherichia coli* (NovaBlue) for initial characterisation. This included colony PCR to confirm that the plasmids carried in each strain contained the correct insert and one of the intended size. The insert DNA was also sequenced from each terminus using vector-specific primers annealing to the S.tag and T7 terminator sequences. Three types of *E. coli* DE3 cells were also evaluated for their capacity to express the recombinant clostridial proteins and the expressed protein products were purified by nickel affinity chromatography for verification by mass spectrometry (LC-MS/MS).

3.2.1 Cloning of the targets

Extracted PCR products for each target were first treated with T4 DNA polymerase to generate the overhanging termini required for LIC cloning, and then they were annealed to the prepared pET-32 vector. Competent *E. coli* NovaBlue cells were used as the strain for initial transformation of the ligation products. Transformations resulted in large numbers of colonies when incubated overnight on selective agar plates, comparable to the positive control insert supplied with the vector kit. Few or no colonies were seen for negative controls. Colonies were picked for colony PCR using target-specific primers (Table 3.1). This showed most transformants carried inserts of the size expected for each target (Figure 3.5). Working from the top of Figure 3.5, it is apparent that 6 out of 10 colonies carrying the *cwp84* sequence were positive in having an insert of the expected size (2265 bp) whereas for *cspA* (189 bp) and 5' *cwp66* (744 bp) all 10 colonies that were picked were positive. The fourth row in the Figure 3.5 shows *fbpA* transformants in

which an insert of 1689 bp was expected. Just one colony from the 10 selected lacked an insert. For the 6 other targets, colony PCR from all transformants showed amplified fragments of the predicted sizes (Figure 3.5).

<p><i>CWP84</i> 2265 BP, 6 OUT OF 10 COLONIES CARRIED INSERTS OF THE CORRECT SIZE.</p>	
<p><i>cspA</i> 189bp, all 10 colonies carried inserts of the correct size.</p>	
<p>5' <i>cwp66</i> 744 bp, all 10 colonies carried inserts of the correct size.</p>	
<p><i>fbpA</i> 1689 bp, 9 out of 10 colonies carried inserts of the correct size.</p>	
<p><i>groEL</i> 1461 bp, all 10 colonies carried inserts of the correct size.</p>	

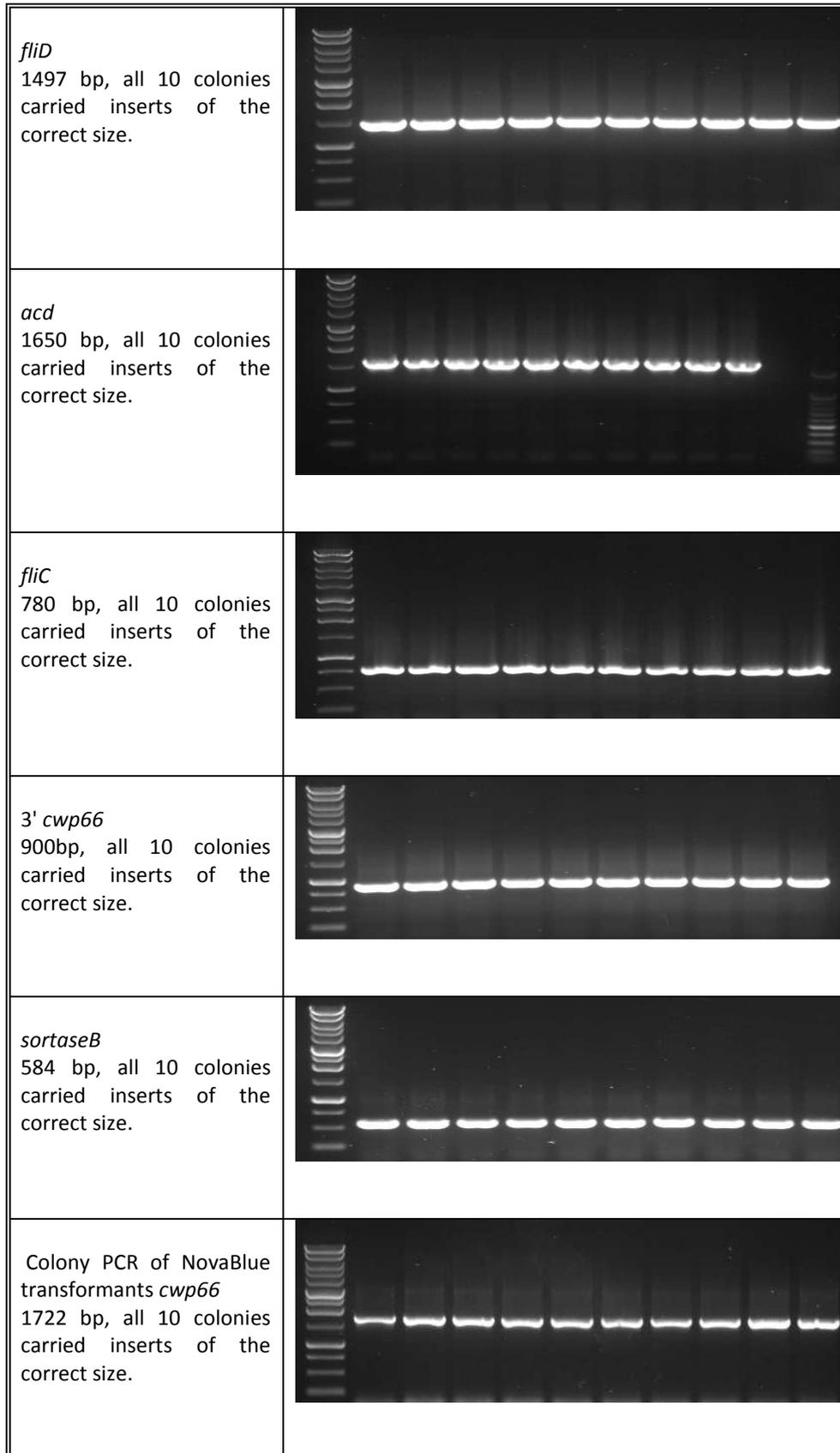


Figure 3. 5. Colony PCR of NovaBlue transformants

lane 1 of all gels, 1kb DNA ladder.

3.2.2 DNA sequencing analysis

The sequence of each insert was determined using S.tag and T7 terminator primers. Based on the sequence of the pET-32 EK/LIC vector system, different restriction endonuclease sites flank the insert. These were all checked in the sequence data collected from each target. In addition, an enterokinase site (EK site), and His.tag sequence were other markers to be checked in the sequences (Table 3.4).

To check inserts which were less than 1000 bp, data collected from S.tag or T7 primed sequencing was compared to ensure that the same results had been obtained and the data was aligned to related DNA sequence from *C. difficile* 630. For inserts more than 1000 bp in length, contigs were assembled from the two sequencing reactions and again compared to data from *C. difficile* 630. The assembled data were translated to their amino acid sequence and applied into BLAST engine (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi). Table 3.4 shows an overview of the cloned sequences with features that were checked such as the EK site, restriction enzyme sites, His.tag and stop codons.

For one clone carrying *cwp84*, sequence analysis showed the termini of the insert were in complete agreement with the reference entry at GenBank, but as shown in Figure 3.6, a central region had been deleted. This construct was designated r-*cwp84*. Comparing three sequences, r-*cwp84* lacked 335 nucleotides in the middle. The cloned sequence for *cspA* was analysed from two S.tag and T7-primer reactions. Data from T7 sequencing was used in an alignment with *cspA630*, the sequence of *cspA* from *C. difficile* 630. The DNA and translated sequences showed completely identity with *cspA630* (Figure 3.8). For the *fbpA*

insert, a fragment of 1689 bp was found in the plasmid which encoded the same protein sequence as *fbpA* of 630 strain (Figure 3.9) and similar results were obtained with *groEL* (1461 bp; Figure 3.10), *acd* (1650 bp; Figure 3.12), and *fliC* (780 bp; Figure 3.13). For *cwp66*, two constructs were generated carrying the 5' (r-5cwp66) and 3' (r-3cwp66) regions of the coding sequence. These were aligned separately with the sequence of *cwp66* from *C. difficile* 630 as can be seen in Figure 3.14. In BLAST analysis, both recombinant targets showed a high identity with the reference sequence (Figure 3.14). In other experiments, the whole sequence of *cwp66* was also cloned and analysis clearly showed identity when compared with the sequence from *C. difficile* 630 (Figure 3.15). The last sequence analysed was a putative sortase (designated r-srB, 579 bp). Both alignment and BLAST results showed the clone had the intended sequence (Figure 3.16). All target sequences were also analysed in nucleotide BLAST. This showed the inserts were of the right length in the expression plasmid except *cwp84* (data not shown; see earlier comments).


```

designed      AGCAGAGATTCTGAAAAA CTGACTCTGTTATGTTGAAACAGATTCTGTAGGAGCAAAA 1035
r-cwp84      AGCAGAGATTCTGAAAAA CTGACTCTGTTATGTTGAAACAGATTCTGTAG----- 1027
630cwp84     AGCAGAGATTCTGAAAAA CTGACTCTGTTATGTTGAAACAGATTCTGTAGGAGCAAAA 1140

designed      TATGAAGTATATTATGCACCAGTAGTAAATGGAGTTCCTCAAAACAATTCAATGACAAAA 1095
r-cwp84      -----
630cwp84     TATGAAGTATATTATGCACCAGTAGTAAATGGAGTTCCTCAAAACAATTCAATGACAAAA 1200

designed      CTTGCAAGTGGAACAGTATCATA TTCTGGATACATAAAATGTACCTACTAATTCTTACAGC 1155
r-cwp84      -----
630cwp84     CTTGCAAGTGGAACAGTATCATA TTCTGGATACATAAAATGTACCTACTAATTCTTACAGC 1260

designed      TTACCAAAAGGTAAGGAGCAATAGTAGTAGTTATAGACAACACAGCAAAATCCTAATAGA 1215
r-cwp84      -----
630cwp84     TTACCAAAAGGTAAGGAGCAATAGTAGTAGTTATAGACAACACAGCAAAATCCTAATAGA 1320

designed      GAAAAATCAACTTTAGCATATGAAACTAATATAGATGCATATTTATTTATATGAGGCTAAA 1275
r-cwp84      -----
630cwp84     GAAAAATCAACTTTAGCATATGAAACTAATATAGATGCATATTTATTTATATGAGGCTAAA 1380

designed      GCAAACCTTAGGTGAAAGCTACATACTTCAAAACAATAAGTTTGAAGACATAAATACATAT 1335
r-cwp84      -----
630cwp84     GCAAACCTTAGGTGAAAGCTACATACTTCAAAACAATAAGTTTGAAGACATAAATACATAT 1440

designed      AGTGAATTTTCTCCTTGTAACTTTGTTATAAAAGCTATAACAAAAACATCTTCTGGACAA 1395
r-cwp84      -----ATAAAAGCTATAACAAAAACATCTTCTGGACAA 1060
630cwp84     AGTGAATTTTCTCCTTGTAACTTTGTTATAAAAGCTATAACAAAAACATCTTCTGGACAA 1500

```

Figure 3.7. Alignment shows the intended sequence (“designed”), that recovered from cloning (“r-cwp84”) and the *cwp84* sequence from *Clostridium difficile* 630 (“630cwp84”).

A

```

cspA630      MKNGIVKWFNNEKGFGFISVEGEDDVFVHFSAIQNDGYKTLEEGEKVSFDITQGNRGPQA
r-cspA      MKNGIVKWFNNEKGFGFISVEGEDDVFVHFSAIQNDGYKTLEEGEKVSFDITQGNRGPQA
*****
cspA360      ENVNRI 66
r-cspA      ENV--- 63

```

B

```

ref|NC_009089.1| D Clostridium difficile 630 chromosome, complete genome
Length=4290252

```

Sort alignments for this subject sequence by:

E value Score Percent identity

Query start position Subject start position

Features in this part of subject sequence:

cold shock protein

Score = 128 bits (322), Expect = 2e-30, Method: Compositional matrix adjust.

Identities = 63/63 (100%), Positives = 63/63 (100%), Gaps = 0/63 (0%)

Frame = -3

```

Query 1      MKNGIVKWFNNEKGFGFISVEGEDDVFVHFSAIQNDGYKTLEEGEKVSFDITQGNRGPQA 60
              MKNGIVKWFNNEKGFGFISVEGEDDVFVHFSAIQNDGYKTLEEGEKVSFDITQGNRGPQA
Sbjct 1074175 MKNGIVKWFNNEKGFGFISVEGEDDVFVHFSAIQNDGYKTLEEGEKVSFDITQGNRGPQA 1073996

Query 61     ENV 63
              ENV
Sbjct 1073995 ENV 1073987

```

Figure 3.8. Alignment of the predicted protein sequence of r-cspA with cspA630 (A) and BLAST output from a search of the protein database (B).

A

```

r-fbpA -----IHQPEDDEVIFNIRNNKENFRLVLSASASNPRVY 34
fbpa630 MALDGLVIHSIVDELSSKLTGGKIDKIHQPEDDEVIFNIRNNKENFRLVLSASASNPRVY 60
          *****

r-fbpA LTSNYQKENPLKAPMFCMLLRKYIQGGNIVEISQIGFERIIKISVESLDELKEKTVKNIM 94
fbpa630 LTSNYQKENPLKAPMFCMLLRKYIQGGNIVEISQIGFERIIKISVESLDELKEKTVKNIM 120
          *****

r-fbpA IEIMGRHSNIIITHGEENKIIDSIKRVVPSISRVRQVLPGHDSLPPQNKLNPLDDISK 154
fbpa630 IEIMGRHSNIIITHGEENKIIDSIKRVVPSISRVRQVLPGHDSLPPQNKLNPLDDISK 180
          *****

r-fbpA DLFIKNLEELGPIFKSIYSKFLGISPIIAKEICYRAGVNQNAI IKDISDEQFDALHKVF 214
fbpa630 DLFIKNLEELGPIFKSIYSKFLGISPIIAKEICYRAGVNQNAI IKDISDEQFDALHKVF 240
          *****

r-fbpA CNLFNDINSNKYSPCIIIDKKVDRVDFSCINLTLFSDLSYINKDSMSRILEDYFRTKDI 274
fbpa630 CNLFNDINSNKYSPCIIIDKKVDRVDFSCINLTLFSDLSYINKDSMSRILEDYFRTKDI 300
          *****

r-fbpA KDRINQRSSDLKKSISVKLDRLYNKLKQEEELSESENADIYKIKGELITSYIYMVEKGM 334
fbpa630 KDRINQRSSDLKKSISVKLDRLYNKLKQEEELSESENADIYKIKGELITSYIYMVEKGM 360
          *****

r-fbpA ESIEVANFYDENCNDVIIELNKNLTPSENAQKYFKKYNKMKHAKVEISHQISLNKEEIDY 394
fbpa630 ESIEVANFYDENCNDVIIELNKNLTPSENAQKYFKKYNKMKHAKVEISHQISLNKEEIDY 420
          *****

r-fbpA LENIILSIENCENLAELQDIKEELAKVGYIKTKKNSKKDTIPSTKPFVSSDGFKILV 454
fbpa630 LENIILSIENCENLAELQDIKEELAKVGYIKTKKNSKKDTIPSTKPFVSSDGFKILV 480
          *****

r-fbpA GKNNKQNDYLTLR LADNDDLWMHTKNIPGSHVIK CAGKEVPDNTVFEGAMLAFFSKSK 514
fbpa630 GKNNKQNDYLTLR LADNDDLWMHTKNIPGSHVIK CAGKEVPDNTVFEGAMLAFFSKSK 540
          *****

r-fbpA LSSQVPVDYTKRKNVKKPSGSKPGMVIYETNSTIYVTPPEETVAKLKVK -- 563
fbpa630 LSSQVPVDYTKRKNVKKPSGSKPGMVIYETNSTIYVTPPEETVAKLKVKSE 591
          *****

```

B  [ref|NC_009089.1|](https://www.ncbi.nlm.nih.gov/blast/)  Clostridium difficile 630 chromosome, complete genome
Length=4290252
Features in this part of subject sequence:
[fibronectin-binding protein](#)
Score = 1108 bits (2867), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 563/563 (100%), Positives = 563/563 (100%), Gaps = 0/563 (0%)
Frame = +3

```

Query 1      IHQPEDDEVIFNIRNNKENFRLVLSASASNPRVYLTSNYQKENPLKAPMFCMLLRKYIQG 60
             IHQPEDDEVIFNIRNNKENFRLVLSASASNPRVYLTSNYQKENPLKAPMFCMLLRKYIQG
Sbjct 2998230 IHQPEDDEVIFNIRNNKENFRLVLSASASNPRVYLTSNYQKENPLKAPMFCMLLRKYIQG 2998409
Query 61      GNIVEISQIGFERI IKISVESLDELKEKTVKNIMIEIMGRHSNIIITHGEENKIIDSIKR 120
             GNIVEISQIGFERI IKISVESLDELKEKTVKNIMIEIMGRHSNIIITHGEENKIIDSIKR
Sbjct 2998410 GNIVEISQIGFERI IKISVESLDELKEKTVKNIMIEIMGRHSNIIITHGEENKIIDSIKR 2998589
Query 121     VPFSSISRVRQVLPGHDSLPPQNKLNPLDDISKDLFIKNLEELGPIFKSIYSKFLGIS 180
             VPFSSISRVRQVLPGHDSLPPQNKLNPLDDISKDLFIKNLEELGPIFKSIYSKFLGIS
Sbjct 2998590 VPFSSISRVRQVLPGHDSLPPQNKLNPLDDISKDLFIKNLEELGPIFKSIYSKFLGIS 2998769
Query 181     PIIAKEICYRAGVNQNAI IKDISDEQFDALHKVFCNLFNDINSNKYSPCIIIDKKVDRVV 240
             PIIAKEICYRAGVNQNAI IKDISDEQFDALHKVFCNLFNDINSNKYSPCIIIDKKVDRVV
Sbjct 2998770 PIIAKEICYRAGVNQNAI IKDISDEQFDALHKVFCNLFNDINSNKYSPCIIIDKKVDRVV 2998949
Query 241     DFSCINLTLFSDLSYINKDSMSRILEDYFRTKDIKDRINQRSSDLKKSISVKLDRLYNKL 300
             DFSCINLTLFSDLSYINKDSMSRILEDYFRTKDIKDRINQRSSDLKKSISVKLDRLYNKL
Sbjct 2998950 DFSCINLTLFSDLSYINKDSMSRILEDYFRTKDIKDRINQRSSDLKKSISVKLDRLYNKL 2999129
Query 301     kKqeeelseSenADIYKIKGELITSYIYMVEKGMESIEVANFYDENCNDVIIELNKNLTP 360
             KKQEEELSESENADIYKIKGELITSYIYMVEKGMESIEVANFYDENCNDVIIELNKNLTP
Sbjct 2999130 KKQEEELSESENADIYKIKGELITSYIYMVEKGMESIEVANFYDENCNDVIIELNKNLTP 2999309
Query 361     SENAQKYFKKYNKMKHAKVEISHQISLNKEEIDYLENIILSIENCENLAELQDIKEELAK 420
             SENAQKYFKKYNKMKHAKVEISHQISLNKEEIDYLENIILSIENCENLAELQDIKEELAK
Sbjct 2999310 SENAQKYFKKYNKMKHAKVEISHQISLNKEEIDYLENIILSIENCENLAELQDIKEELAK 2999489
Query 421     VGYIKTQKNSKKDTIPSTKPFVSSDGFKILVGKNNKQNDYLTLR LADNDDLWMHTKN 480
             VGYIKTQKNSKKDTIPSTKPFVSSDGFKILVGKNNKQNDYLTLR LADNDDLWMHTKN
Sbjct 2999490 VGYIKTQKNSKKDTIPSTKPFVSSDGFKILVGKNNKQNDYLTLR LADNDDLWMHTKN 2999669
Query 481     IPGSHVIK CAGKEVPDNTVFEGAMLAFFSKSKLSSQVPVDYTKRKNVKKPSGSKPGMV 540
             IPGSHVIK CAGKEVPDNTVFEGAMLAFFSKSKLSSQVPVDYTKRKNVKKPSGSKPGMV
Sbjct 2999670 IPGSHVIK CAGKEVPDNTVFEGAMLAFFSKSKLSSQVPVDYTKRKNVKKPSGSKPGMV 2999849
Query 541     IYETNSTIYVTPPEETVAKLKVK 563
             IYETNSTIYVTPPEETVAKLKVK
Sbjct 2999850 IYETNSTIYVTPPEETVAKLKVK 2999918

```

Figure 3. 9. Alignment of the predicted protein sequence of r-fbpA with fbpa630 (A) and BLAST output from a search of the protein database (B).

A

```

groEL630      MAKEIKFSEETRRALEAGVNKLADTVKVTLPKGRNVILDKFKGSPPLITNDGVTIAKEIE 60
r-groEL      -----IGVTIAKEIE 10
                                     *****

groEL630      LEDRFENMGAQLVKEVATKTNDVAGDGTTTATVLAQAIIREGLKNVTAGANPILLRKGIQ 120
r-groEL      LEDRFENMGAQLVKEVATKTNDVAGDGTTTATVLAQAIIREGLKNVTAGANPILLRKGIQ 70
                                     *****

groEL630      KAVTVAVEELKNQSRIVETQEAISQVASISAGDEEVGKLIAEAMEIVGKDGVITVEESQT 180
r-groEL      KAVTVAVEELKNQSRIVETQEAISQVASISAGDEEVGKLIAEAMEIVGKDGVITVEESQT 130
                                     *****

groEL630      MNTELDAVEGMQFDRGFVSAYMVTDVDKMEAVLNDPYILITDKKISNIQELLPVLEQIVQ 240
r-groEL      MNTELDAVEGMQFDRGFVSAYMVTDVDKMEAVLNDPYILITDKKISNIQELLPVLEQIVQ 190
                                     *****

groEL630      QGKLLIIAEDVEGEALSTLVVNKLRGTFDVVAVKAPGFGDRRKEMLQDIAILTGAQVIS 300
r-groEL      QGKLLIIAEDVEGEALSTLVVNKLRGTFDVVAVKAPGFGDRRKEMLQDIAILTGAQVIS 250
                                     *****

groEL630      EELGYDLKEADLSMLGRASSVKVTKESTTIVDGSGDKKAIEDRVTQIKHQVEQTTSDFDR 360
r-groEL      EELGYDLKEADLSMLGRASSVKVTKESTTIVDGSGDKKAIEDRVTQIKHQVEQTTSDFDR 310
                                     *****

groEL630      EKLMERLAKLAGGVAVVKVGAATEVELKERKLRIEDALNATRAAVEEGIVAGGGTAFVSV 420
r-groEL      EKLMERLAKLAGGVAVVKVGAATEVELKERKLRIEDALNATRAAVEEGIVAGGGTAFVSV 370
                                     *****

groEL630      IPAIGTLIESLEGEVKLGAQIVKKALEELRQIAINAGLEGAVIQNVVNSEAETGFDAL 480
r-groEL      IPAIGTLIESLEGEVKLGAQIVKKALEELRQIAINAGLEGAVIQNVVNSEAETGFDAL 430
                                     *****

groEL630      NEKYVNMIEAGIVDPTKVRSALQNAASIASTFLTTEAAVADLPEKEDAGMPGMGGGMPG 540
r-groEL      NEKYVNMIEAGIVDPTKVRSALQNAASIASTFLTTEAAVADLPEKEDAGMPGMGGGMPG --- 487
                                     *****

groEL630      MM 542
r-groEL      --

```

B

> [ref|NC_009089.1](#)  Clostridium difficile 630 chromosome, complete genome Length=4290252 Features
in this part of subject sequence:
60 kDa chaperonin
Score = 863 bits (2230), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 478/478 (100%), Positives = 478/478 (100%), Gaps = 0/478 (0%)
Frame = +2

```

Query 2      GVTIAKEIELEDRFENMGAQLVKEVATKTNDVAGDGTTTATVLAQAIIREGLKNVTAGAN 61
Sbjct 252413 GVTIAKEIELEDRFENMGAQLVKEVATKTNDVAGDGTTTATVLAQAIIREGLKNVTAGAN 252592

Query 62     PILLRKGIQKAVTVAVEELKNQSRIVETQEAISQVASISAGDEEVGKLIAEAMEIVGKDG 121
Sbjct 252593 PILLRKGIQKAVTVAVEELKNQSRIVETQEAISQVASISAGDEEVGKLIAEAMEIVGKDG 252772

Query 122    VITVEESQTMNTELDAVEGMQFDRGFVSAYMVTDVDKMEAVLNDPYILITDKKISNIQEL 181
Sbjct 252773 VITVEESQTMNTELDAVEGMQFDRGFVSAYMVTDVDKMEAVLNDPYILITDKKISNIQEL 252952

Query 182    LPVLEQIVQGGKLLIIAEDVEGEALSTLVVNKLRGTFDVVAVKAPGFGDRRKEMLQDIA 241
Sbjct 252953 LPVLEQIVQGGKLLIIAEDVEGEALSTLVVNKLRGTFDVVAVKAPGFGDRRKEMLQDIA 253132

Query 242    ILTGAQVISEELGYDLKEADLSMLGRASSVKVTKESTTIVDGSGDKKAIEDRVTQIKHQV 301
Sbjct 253133 ILTGAQVISEELGYDLKEADLSMLGRASSVKVTKESTTIVDGSGDKKAIEDRVTQIKHQV 253312

Query 302    EQTTSDFDREKLMERLaklaggvavvkvgaaTEVELKERKLRIEDALNATRAAVEEGIVA 361
Sbjct 253313 EQTTSDFDREKLMERLaklaggvavvkvgaaTEVELKERKLRIEDALNATRAAVEEGIVA 253492

Query 362    gggtaFVSVIPAIGTLIESLEGEVKLGAQIVKKALEELRQIAINAGLEGAVIQNVVNS 421
Sbjct 253493 gggtaFVSVIPAIGTLIESLEGEVKLGAQIVKKALEELRQIAINAGLEGAVIQNVVNS 253672

Query 422    EAETGFDALNEKYVNMIEAGIVDPTKVRSALQNAASIASTFLTTEAAVADLPEKEDA 479
Sbjct 253673 EAETGFDALNEKYVNMIEAGIVDPTKVRSALQNAASIASTFLTTEAAVADLPEKEDA 253846

```

Figure 3. 10 Alignment of the predicted protein sequence of r-groEL with groEL630 (A) and BLAST output from a search of the protein database (B).

A

```

r-fliD      ----IPVVRTGLSGNFDMEGIIIEASMIIRDKEKVNKAKQDQQIVKWKQEIYRDIIEKSKNL 56
fliD630     MSSISPVVRTGLSGNFDMEGIIIEASMIIRDKEKVNKAKQDQQIVKWKQEIYRDIIEKSKNL 60
            *****

r-fliD      YDKYLNQDSPNSITNKKAYSATRITSSDESIIIVAKGSAGA EKINYQFAVSQMAEPAKVTI 116
fliD630     YDKYLNQDSPNSITNKKAYSATRITSSDESIIIVAKGSAGA EKINYQFAVSQMAEPAKVTI 120
            *****

r-fliD      KLNSSDPIVQQFPPNAGASSLNIGGVNIPISEQD TTSTIVSKINSLCADNDIRASYSEM 176
fliD630     KLNSSDPIVQQFPPNAGASSLNIGGVNIPISEQD TTSTIVSKINSLCADNDIRASYSEM 180
            *****

r-fliD      TGEIIISRKQTGSSSDIDLKVI GNDSLAGQIASDNGITFTT DASGTKSAVVY GKNLEADV 236
fliD630     TGEIIISRKQTGSSSDIDLKVI GNDSLAGQIASDNGITFTT DASGTKSAVVY GKNLEADV 240
            *****

r-fliD      TDDQGRVTHISKEQNSFKIDNIDYVNSKGS AKLVSVDTEEATKNMKAFVDVYNALMDK 296
fliD630     TDDQGRVTHISKEQNSFKIDNIDYVNSKGS AKLVSVDTEEATKNMKAFVDDYNALMDK 300
            *****

r-fliD      VYGLVTTKKS KDYPPLTDEQKDDMTTEEIEKWEKKAKEGILRNDELRAFVEDIQSMFFG 356
fliD630     VYGLVTTKKS KDYPPLTDEQKDDMTTEEIEKWEKKAKEGILRNDELRAFVEDIQSMFFG 360
            *****

r-fliD      DADTIIALRKLGI SEHENYKKGQISFNADTF SKALIDSDKVYKALAGYSSNYDDKGMF 416
fliD630     DADTIIALRKLGI SEHENYKKGQISFNADTF SKALIDSDKVYKALAGYSSNYDDKGMF 420
            *****

r-fliD      EKLK KIVFEYSGSSASKLTKKAGMENS SSSASQNVYSKQIAEQERNISRLVEKMNDKEKRL 476
fliD630     EKLK KIVFEYSGSSASKLTKKAGMENS SSSASQNVYSKQIAEQERNISRLVEKMNDKEKRL 480
            *****

r-fliD      YAKYSALESLLNKYSSQMNYFSQ---- 499
fliD630     YAKYSALESLLNKYSSQMNYFSQAQGN 507

```

B > [ref|NC_009089.1](#)  Clostridium difficile 630 chromosome, complete genome
Length=4290252
Features in this part of subject sequence:
[flagellar cap protein](#)
Score = 974 bits (2519), Expect = 0.0, Method: Compositional matrix adjust. Identities = 497/498 (99%), Positives = 497/498 (99%), Gaps = 0/498 (0%)
Frame = +3

```

Query 2      PVRVTGLSGNFDMEGIIIEASMIIRDKEKVNKAKQDQQIVKWKQEIYRDIIEKSKNL YDKYL 61
Sbjct 298995 PVRVTGLSGNFDMEGIIIEASMIIRDKEKVNKAKQDQQIVKWKQEIYRDIIEKSKNL YDKYL 299174
Query 62     NGDSPNSITNKKAYSATRITSSDESIIIVAKGSAGA EKINYQFAVSQMAEPAKVTIKLNSS 121
Sbjct 299175 NGDSPNSITNKKAYSATRITSSDESIIIVAKGSAGA EKINYQFAVSQMAEPAKVTIKLNSS 299354
Query 122    DPIVQQFPPNAGASSLNIGGVNIPISEQD TTSTIVSKINSLCADNDIRASYSEM TGEI 181
Sbjct 299355 DPIVQQFPPNAGASSLNIGGVNIPISEQD TTSTIVSKINSLCADNDIRASYSEM TGEI 299534
Query 182    ISRKQTGSSSDIDLKVI GNDSLAGQIASDNGITFTT DASGTKSAVVY GKNLEADV TDDQ 241
Sbjct 299535 ISRKQTGSSSDIDLKVI GNDSLAGQIASDNGITFTT DASGTKSAVVY GKNLEADV TDDQ 299714
Query 242    RVTHISKEQNSFKIDNIDYVNSKGS AKLVSVDTEEATKNMKAFVDVYNALMDKVYGLV 301
Sbjct 299715 RVTHISKEQNSFKIDNIDYVNSKGS AKLVSVDTEEATKNMKAFVDVYNALMDKVYGLV 299894
Query 302    TTKKSKDYPPLTDEQKDDMTTEEIEKWEKKAKEGILRNDELRAFVEDIQSMFFGDADTI 361
Sbjct 299895 TTKKSKDYPPLTDEQKDDMTTEEIEKWEKKAKEGILRNDELRAFVEDIQSMFFGDADTI 300074
Query 362    IALRKLGI SEHENYKKGQISFNADTF SKALIDSDKVYKALAGYSSNYDDKGMFEKLLK 421
Sbjct 300075 IALRKLGI SEHENYKKGQISFNADTF SKALIDSDKVYKALAGYSSNYDDKGMFEKLLK 300254
Query 422    IVFEYSGSSASKLTKKAGMENS SSSASQNVYSKQIAEQERNISRLVEKMNDKEKRL YAKYS 481
Sbjct 300255 IVFEYSGSSASKLTKKAGMENS SSSASQNVYSKQIAEQERNISRLVEKMNDKEKRL YAKYS 300434
Query 482    ALESLLNKYSSQMNYFSQ 499
Sbjct 300435 ALESLLNKYSSQMNYFSQ 300488

```

Figure 3. 11. Alignment of the predicted protein sequence of r-fliD with fliD630 (A) and BLAST output from a search of the protein database (B).

A

```

acd630      MKKKAALATLAMLPLGVVNAHADGDIGIVTINYLNVRNEPTAESSIAFVAKKDDKVLIKD 60
r-acd      -----I-EPTAESSIAFVAKKDDKVLIKD 23
          *****

acd630      SSNGWYKIKAESGQEGWASSKYIAKSNSDSLRTSTNKEKQVISNSLNMNRNGAGTSYRVIT 120
r-acd      SSNGWYKIKAESGQEGWASSKYIAKSNSDSLRTSTNKEKQVISNSLNMNRNGAGTSYRVIT 83
          *****

acd630      VLKKGQKVEVISESNGWSKIKYDGRGLGYVSSSYLGDVSNSTNKSKTKQVNTTSLNVRSGP 180
r-acd      VLKKGQKVEVISESNGWSKIKYDGRGLGYVSSSYLGDVSNSTNKSKTKQVNTTSLNVRSGP 143
          *****

acd630      NTSYGLLGKLPKGSKVEVISESNGWSKIKYNGKDAYVSSMYLSDVQSNSDNSSQSNDKK 240
r-acd      NTSYGLLGKLPKGSKVEVISESNGWSKIKYNGKDAYVSSMYLSDVQSNSDNSSQSNDKK 203
          *****

acd630      NTDKVVNTASLNVRSGPGSTYSKLGKVKYKGSKVTVLSESSGWAKINFNNKEAFVVGNYLS 300
r-acd      NTDKVVNTASLNVRSGPGSTYSKLGKVKYKGSKVTVLSESSGWAKINFNNKEAFVVGNYLS 263
          *****

acd630      TSADTSNNSNSNYDNSSNSNGNNSSSSQVNGMSGISGAKIDYKSLSYTLESHISKQVE 360
r-acd      TSADTSNNSNSNYDNSSNSNGNNSSSSQVNGMSGISGAKIDYKSLSYTLESHISKQVE 323
          *****

acd630      KAASGGNVIAPSNRKSTPSEFSTFSAQRTSSFVNASSSDIEYYLNPKNFNTTTKGMMQF 420
r-acd      KAASRGNVIAPSNRKSTPSEFSTFSAQRTSSFVNASSSDIEYYLNPKNFNTTTKGMMQF 383
          *****

acd630      LKINSYRDGISESSLNSYLNGLSSSVFKNQGAAFINAACKYINIDVVYLVSHAMWETAYGK 480
r-acd      LKINSYRDGISESSLNSYLNGLSSSVFKNQGAAFINAACKYINIDVVYLVSHAMWETAYGK 443
          *****

acd630      STLAQQQTLTYSYKQGPLSKPVKYVNFYFGAIDK SANVSGAEAAYSNGWTSVETTIDGSA 540
r-acd      STLAQQQTLTYSYKQGPLSKPVKYVNFYFGAIDK SANVSGAEAAYSNGWTSVETTIDGSA 503
          *****

acd630      KWI SQNYVNSSKYNQNTIYKMKWNYDYTWHQYATDVNWANGISGIMENLIGLYRGSSSLV 600
r-acd      KWI SQNYVNSSKYNQNTIYKMKWNYDYTWHQYATDVNWANGISGIME----- 550
          *****

acd630      FEVPQYK 607
r-acd      -----

```

B  ref|NC_009089.1|  Clostridium difficile 630 chromosome, complete genome
Length=4290252
Sort alignments for this subject sequence by:
E value Score Percent identity
Query start position Subject start position
Features in this part of subject sequence:
putative mannosyl-glycoprotein endo-beta-N-acetylglucosam...

Score = 961 bits (2484), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 547/549 (99%), Positives = 547/549 (99%), Gaps = 0/549 (0%)
Frame = +1

```

Query 2      EPTAESSIAFVAKKDDKVLIKDSSNGWYKIKAESGQEGWASSKYIAKSNSDSLRTSTNKE 61
Sbjct 1509841 EPTAESSIAFVAKKDDKVLIKDSSNGWYKIKAESGQEGWASSKYIAKSNSDSLRTSTNKE 1510020
Query 62      QVISNSLNMNRNGAGTSYRVITVLKKGQKVEVISESNGWSKIKYDGRGLGYVSSSYLGDVS 121
Sbjct 1510021 QVISNSLNMNRNGAGTSYRVITVLKKGQKVEVISESNGWSKIKYDGRGLGYVSSSYLGDVS 1510200
Query 122     NSTNKSCTKQVNTTSLNVRSGPNTSYGLLGKLPKGSKVEVISESNGWSKIKYNGKDAYVS 181
Sbjct 1510201 NSTNKSCTKQVNTTSLNVRSGPNTSYGLLGKLPKGSKVEVISESNGWSKIKYNGKDAYVS 1510380
Query 182     SMYLSDVQSNSDNSSQSNDKKNNTDKVVNTASLNVRSGPGSTYSKLGKVKYKGSKVTVLSE 241
Sbjct 1510381 SMYLSDVQSNSDNSSQSNDKKNNTDKVVNTASLNVRSGPGSTYSKLGKVKYKGSKVTVLSE 1510560
Query 242     SSGWAKINFNNKEAFVVGNYLSTADTSNNSNSNYDNSSNSNGNNSSSSQVNGMSGIS 301
Sbjct 1510561 SSGWAKINFNNKEAFVVGNYLSTADTSNNSNSNYDNSSNSNGNNSSSSQVNGMSGIS 1510740
Query 302     GAKIDYKSLSYTLESHISKQVEKAASRGNVIAPSNRKSTPSEFSTFSAQRTSSFVNASS 361
Sbjct 1510741 GAKIDYKSLSYTLESHISKQVEKAASGGNVIAPSNRKSTPSEFSTFSAQRTSSFVNASS 1510920
Query 362     SDIEYYLNPKNFNTTTKGMMQFLKINSYRDGISESSLNSYLNGLSSSVFKNQGAAFINAA 421
Sbjct 1510921 SDIEYYLNPKNFNTTTKGMMQFLKINSYRDGISESSLNSYLNGLSSSVFKNQGAAFINAA 1511100
Query 422     KKYINIDVVYLVSHAMWETAYGKSTLAQQQTLTYSYKQGPLSKPVKYVNFYFGAIDK SANV 481
Sbjct 1511101 KKYINIDVVYLVSHAMWETAYGKSTLAQQQTLTYSYKQGPLSKPVKYVNFYFGAIDK SANV 1511280
Query 482     SGAEAAYSNGWTSVETTIDGSAKWI SQNYVNSSKYNQNTIYKMKWNYDYTWHQYATDVNW 541
Sbjct 1511281 SGAEAAYSNGWTSVETTIDGSAKWI SQNYVNSSKYNQNTIYKMKWNYDYTWHQYATDVNW 1511460
Query 542     ANGISGIME 550
Sbjct 1511461 ANGISGIME 1511487

```

Figure 3. 12. Alignment of the predicted protein sequence of r-acd with acd630 (A) and BLAST output from a search of the protein database (B).

A

```

flic630      MRVNTNVSALIANNQMGFRNVNGQSKSMEKLSSGVRIKRAADDAAGLAISEKMRAQIKGLD 60
r-flic      -----MEKLSSGVRIKRAADDAAGLAISEKMRAQIKGLD 34
                *****

flic630      QAGRNVQDGISVVQTAEGSLEETGNI LQRMRTLSLQSANEINNTEEREKIADELTLQKDE 120
r-flic      QAGRNVQDGISVVQTAEGSLEETGNI LQRMRTLSLQSANEINNTEEREKIADELTLQKDE 94
                *****

flic630      IERISSSTEFNGKLLDGTSSSTIRLQVGASYGTNVSGTSNNNNEIKIQLVNTASIMASAG 180
r-flic      IERISSSTEFNGKLLDGTSSSTIRLQVGASYGTNVSGTSNNNNEIKIQLVNTASIMASAG 154
                *****

flic630      ITTASIGSMKAGGTTGTDAAKTMVSSLDAAKSLNSSRAKLGAAQNRLSTQNNLNNTLE 240
r-flic      ITTASIGSMKAGGTTGTDAAKTMVSSLDAAKSLNSSRAKLGAAQNRLSTQNNLNNTLE 214
                *****

flic630      NVTAAESRIRDVDVASEMVNLSKMNILVQASQSMLAQANQQPQGV LQLL 290
r-flic      NVTAAESRIRDVDVASEMVNLSKMNILVQASQSMLAQANQQPQGV L--- 260
                *****

```

B   ref|NC_009089.1| Clostridium difficile 630 chromosome, complete genome
Length=4290252
Features in this part of subject sequence:
flagellin subunit
Score = 457 bits (1175), Expect = 1e-128, Method: Compositional matrix adjust.
Identities = 260/260 (100%), Positives = 260/260 (100%), Gaps = 0/260 (0%)
Frame = +3

```

Query 1      MEKLSSGVRIKRAADDAAGLAISEKMRAQIKGLDQAGRNVQDGISVVQTAEGSLEETGNI 60
Sbjct 301020  MEKLSSGVRIKRAADDAAGLAISEKMRAQIKGLDQAGRNVQDGISVVQTAEGSLEETGNI 301199

Query 1      LQRMRTLSLQSANEINNTEEREKIADELTLQKDEIERISSSTEFNGKLLDGTSSSTIRLQ 120
Sbjct 301200  LQRMRTLSLQSANEINNTEEREKIADELTLQKDEIERISSSTEFNGKLLDGTSSSTIRLQ 301379

Query 121    VGASYgtnvsgtsnnnnEIKIQLVNTASIMASAGITTASIGSMKaggttgtdaakTMVSS 180
Sbjct 301380  VGASYGTNVSGTSNNNNEIKIQLVNTASIMASAGITTASIGSMKAGGTTGTDAAKTMVSS 301559

Query 181    LDAALKSLNSSRAKLGAAQNRLSTQNNLNNTLENVTAAESRIRDVDVASEMVNLSKMNI 240
Sbjct 301560  LDAALKSLNSSRAKLGAAQNRLSTQNNLNNTLENVTAAESRIRDVDVASEMVNLSKMNI 301739

Query 241    LVQASQSMLAQANQQPQGV L 260
Sbjct 301740  LVQASQSMLAQANQQPQGV L 3017996

```

Figure 3. 13. Alignment of the predicted protein sequence of r-fliC with fliC630 (A) and BLAST output from a search of the protein database (B).

A

```

cwp66-630      MKISKKIVSLLTMTFLTTLVLYGNTSNASTKDTLLTGSGRWETAIKISQAGWTKSES AVLVN 60
r-5cwp66      -----ITGSGRWETAIKISQAGWTKSES AVLVN 28
                : *****
cwp66-630      DNSIADALSATPFAKAKDAPILLTQSNKLDSTRKAEKRLGVKNVYLIGGSIALSSEIEK 120
r-5cwp66      DNSIADALSATPFAKAKDAPILLTQSNKLDSTRKAEKRLGVKNVYLIGGSIALSSEIEK 88
                *****
cwp66-630      QLNAENINFERISGNSRYDTSCLKLAEKLDREKSIKIVVNGEKLADAVSVGAIAAQEN 180
r-5cwp66      QLNAENINFERISGNSRYDTSCLKLAEKLDREKSIKIVVNGEKLADAVSVGAIAAQEN 148
                *****
cwp66-630      MPIILSDSENGTEVADNFIDSKDIKSYVIGGYSISNSVERSLPNATRIAGSSRSETNA 240
r-5cwp66      MPIILSDSENGTEVADNFIDSKDIKSYVIGGYSISNSVERSLPNATRIAGSSRSETNA 208
                *****
cwp66-630      KIEEFYKDTDIKNIYVTKDGTKNKNDLIDSLAVGVLAAKNSSPIVLAGNKLDTTQKDV 300
r-5cwp66      KIEEFYKDTDIKNIYVTKDGTKNKNDLIDSLAVGVLAAK----- 248
                *****
cwp66-630      NTKIIDKVTQIGGLGNENVEDILDIQEETKYTVETIDELNAAIKRADANDIIKFKPEKE 360
r-3cwp66      -----IVTQIGGLGNENVEDILDIQEETKYTVETIDELNAAIKRADANDIIKFKPEKE 54
                *****
cwp66-630      KTINNSFSIETKKTVTIELDGRYQTITLDIPNGKFNNYAEIEGGVKLNKIKNESLVNKG 420
r-3cwp66      KTINNSFSIETKKTVTIELDGRYQTITLDIPNGKFNNYAEIEGGVKLNKIKNESLVNKG 114
                *****
cwp66-630      SIQDLDIYDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGN 480
r-3cwp66      SIQDLDIYDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGN 174
                *****
cwp66-630      IDVTGKKEPAISGNKPKVNDTEKETKAARGLNPRVEACSVPKDYVMITIPNSPKDSRY 540
r-3cwp66      IDVTGKKEPAISGNKPKVNDTEKETKAARGLNPRVEACSVPKDYVMITIPNSPKDSRY 234
                *****
cwp66-630      KIYRVRVYNKPYAMDVGDKINIGEWTVAPTDEEPPLEKAKNGCYVEAVEVNTSTKEVSRW 600
r-3cwp66      KIYRVRVYNKPYAMDVGDKINIGEWTVAPTDEEPPLEKAKNGCYVEAVEVNTSTKEVSRW 294
                *****
cwp66-630      GRTNATDDGF 610
r-3cwp66      GRTNAT---- 300
                *****
    
```

B

[ref|NC_009089.1|](#) **D** Clostridium difficile 630 chromosome, complete genome

Length=4290252

Sort	alignments	for	this	subject	sequence	by:E	value	Score	Percent	identity
Features in this part of subject sequence:										
cell surface protein										
Score = 485 bits (1248), Expect = 5e-137, Method: Compositional matrix adjust.										
Identities = 247/248 (99%), Positives = 248/248 (100%), Gaps = 0/248 (0%)										
Frame = -2										
Query	1				ITGSGRWETAIKISQAGWTKSES AVLVNDNSIADALSATPFAKAKDAPILLTQSNKLDSTR			60		
Sbjct	3246740				LTGSGRWETAIKISQAGWTKSES AVLVNDNSIADALSATPFAKAKDAPILLTQSNKLDSTR			3246561		
Query	61				TKAELKRLGVKNVYLIGGSIALSSEIEKQLNAENINFERISGNSRYDTSCLKLAEKLDREK			120		
Sbjct	3246560				TKAELKRLGVKNVYLIGGSIALSSEIEKQLNAENINFERISGNSRYDTSCLKLAEKLDREK			3246381		
Query	121				SISKIVVNGEKLADAVSVGAIAAQENMPIILSDSENGTEVADNFIDSKDIKSYVIGG			180		
Sbjct	3246380				SISKIVVNGEKLADAVSVGAIAAQENMPIILSDSENGTEVADNFIDSKDIKSYVIGG			3246201		
Query	181				TYSISNSVERSLPNATRIAGSSRSETNAKIEEFYKDTDIKNIYVTKDGTKNKNDLIDSL			240		
Sbjct	3246200				TYSISNSVERSLPNATRIAGSSRSETNAKIEEFYKDTDIKNIYVTKDGTKNKNDLIDSL			3246021		
Query	241				AVGVLAAK 248					
Sbjct	3246020				AVGVLAAK 3245997					

```

C>   ref|NC_009089.1| Clostridium difficile 630 chromosome, complete genome
Length=4290252 Sort alignments for this subject sequence by:
value Score Percent identity Query start position Subject start position
Features in this part of subject sequence:
  cell surface protein
Score = 568 bits (1463), Expect = 8e-162, Method: Compositional matrix adjust.
Identities = 299/299 (100%), Positives = 299/299 (100%), Gaps = 0/299 (0%)
Frame = -2
Query  2          VTQIGGLGNENVVEDILDIQEETKYTVETIDELNAAIKRADANDIIKFKPEKEKTINNSF 61
                VTQIGGLGNENVVEDILDIQEETKYTVETIDELNAAIKRADANDIIKFKPEKEKTINNSF
Sbjct  3245915 VTQIGGLGNENVVEDILDIQEETKYTVETIDELNAAIKRADANDIIKFKPEKEKTINNSF 3245736
Query  62          SIETKKTVTIELDGRYRQTITLDIPNGKFNNYAEIEGGVKLKNIKNESLVNKGSIQDLDI 121
                SIETKKTVTIELDGRYRQTITLDIPNGKFNNYAEIEGGVKLKNIKNESLVNKGSIQDLDI
Sbjct  3245735 SIETKKTVTIELDGRYRQTITLDIPNGKFNNYAEIEGGVKLKNIKNESLVNKGSIQDLDI 3245556
Query  122         YDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGNIDVTGK 181
                YDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGNIDVTGK
Sbjct  3245555 YDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGNIDVTGK 3245376
Query  182         KEPAISGNKPKVNDTEKETKAARGLNPRVEACSVPKKDYVMITIPNSPKDSRYKIYYRVV 241
                KEPAISGNKPKVNDTEKETKAARGLNPRVEACSVPKKDYVMITIPNSPKDSRYKIYYRVV
Sbjct  3245375 KEPAISGNKPKVNDTEKETKAARGLNPRVEACSVPKKDYVMITIPNSPKDSRYKIYYRVV 3245196
Query  242         YNKPYAMDVGDKINIGEWTVAPTDEEPFLEKAKNGCYVEAVEVNTSTKEVSRWGRTNAT 300
                YNKPYAMDVGDKINIGEWTVAPTDEEPFLEKAKNGCYVEAVEVNTSTKEVSRWGRTNAT
Sbjct  3245195 YNKPYAMDVGDKINIGEWTVAPTDEEPFLEKAKNGCYVEAVEVNTSTKEVSRWGRTNAT 3245019

```

Figure 3. 14. Alignment of the predicted protein sequence of r-5cwp66 and r-3cwp66 with cwp66 630 (A) and BLAST output from a search of the protein database (B and C).

A

```

cwp66-630      MKISKKIVSLLTMTFLTVLYGNTSNASTKDTLTGSGRWETAIKISQAGWTKSESAVLVN 60
r-cwp66      -----ITGSGRWETAIKISQAGWTKSESAVLVN 28
                : *****

cwp66-630      DNSIADALSATPFAKAKDAPILLTQSNKLDSTRKAELKRLGVKNVYLIIGGSIALSSEIEK 120
r-cwp66      DNSIADALSATPFAKAKDAPILLTQSNKLDSTRKAELKRLGVKNVYLIIGGSIALSSEIEK 88
                *****

cwp66-630      QLNNAENINFERISGNSRYDTSCLKLAEKLDREKSIKIVVVNGEKGADAVSVGAIAAQEN 180
r-cwp66      QLNNAENINFERISGNSRYDTSCLKLAEKLDREKSIKIVVVNGEKGADAVSVGAIAAQEN 148
                *****

cwp66-630      MPIILSDSENGTEVADNFIDSKDIAKSYVIGGTYISNSVERSLPNATRIAGSSRSETNA 240
r-cwp66      MPIILSDSENGTEVADNFIDSKDIAKSYVIGGTYISNSVERSLPNATRIAGSSRSETNA 208
                *****

cwp66-630      KIIIEFYKDTDIKNIYVTKDGTKNKNDLIDSLAVGVLAAKNSSPIVLAGNKLDTTQKDV 300
r-cwp66      KIIIEFYKDTDIKNIYVTKDGTKNKNDLIDSLAVGVLAAKNSSPIVLAGNKLDTTQKDV 268
                *****

cwp66-630      NTKIIDKVTQIGGLGNENVVEDILDIEETKYTVETIDELNAAIKRADANDIIKFKPEKE 360
r-cwp66      NTKIIDKVTQIGGLGNENVVEDILDIEETKYTVETIDELNAAIKRADANDIIKFKPEKE 328
                *****

cwp66-630      KTINNSFSIETKKTVTIELDGRYRQTITLDIPNGKFNNYAEIEGGVKLNKIKNESLVNKG 420
r-cwp66      KTINNSFSIETKKTVTIELDGRYRQTITLDIPNGKFNNYAEIEGGVKLNKIKNESLVNKG 388
                *****

cwp66-630      SIQDLDIYDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGN 480
r-cwp66      SIQDLDIYDENGCKIENESSGEIWFVTIVEEANDVYIVNSGDITKISNNSSTIIRNSGN 448
                *****

cwp66-630      IDVTGKKEPAISGNPKPVNDTEKETKAARGLNPRVEACSVPKDYVMITIPNSPKDSRY 540
r-cwp66      IDVTGKKEPAISGNPKPVNDTEKETKAARGLNPRVEACSVPKDYVMITIPNSPKDSRY 508
                *****

cwp66-630      KIYRVVYNKPYAMDVGDKINIGEWTVAPTDEEPEFLEKAKNGCYVEAVEVNTSTKEVSRW 600
r-cwp66      KIYRVVYNKPYAMDVGDKINIGEWTVAPTDEEPEFLEKAKNGCYVEAVEVNTSTKEVSRW 568
                *****

cwp66-630      GRTNATDDGF 610
r-cwp66      GRTNAT---- 574
    
```

B  ref|NC_009089.1 **D** Clostridium difficile 630 chromosome, complete genome

Length=4290252 Sort alignments for this subject sequence by:

E value [Score](#) [Percent identity](#) [Query start position](#) [Subject start position](#)

Features in this part of subject sequence: [cell surface protein](#)

Score = 1104 bits (2856), Expect = 0.0, Method: Compositional matrix adjust.

Identities = 573/574 (99%), Positives = 574/574 (100%), Gaps = 0/574 (0%)

Frame = -2

```

Query 1      ITGSGRWETAIKISQAGWTKSESAVLVNDNSIADALSATPFAKAKDAPILLTQSNKLD 60
Sbjct 3246740 +ITGSGRWETAIKISQAGWTKSESAVLVNDNSIADALSATPFAKAKDAPILLTQSNKLD 3246561
Query 61      TKAEKRLGVKNVYLIIGGSIALSSEIEKQLNAENINFERISGNSRYDTSCLKLAEKLDREK 120
Sbjct 3246560 TKAEKRLGVKNVYLIIGGSIALSSEIEKQLNAENINFERISGNSRYDTSCLKLAEKLDREK 3246381
Query 121     SISKIVVVNGEKGADAVSVGAIAAQENMPIILSDSENGTEVADNFIDSKDIAKSYVIGG 180
Sbjct 3246380 SISKIVVVNGEKGADAVSVGAIAAQENMPIILSDSENGTEVADNFIDSKDIAKSYVIGG 3246201
Query 181     TYSISNSVERSLPNATRIAGSSRSETNAKIIIEEFYKDTDIKNIYVTKDGTKNKNDLIDSL 240
Sbjct 3246200 TYSISNSVERSLPNATRIAGSSRSETNAKIIIEEFYKDTDIKNIYVTKDGTKNKNDLIDSL 3246021
Query 241     AVGVLAAKNSSPIVLAGNKLDTTQKDVLTNTKIIDKVTQIGGLGNENVVEDILDIEETKY 300
Sbjct 3246020 AVGVLAAKNSSPIVLAGNKLDTTQKDVLTNTKIIDKVTQIGGLGNENVVEDILDIEETKY 3245841
Query 301     TVETIDELNAAIKRADANDIIKFKPEKEKTINNSFSIETKKTVTIELDGRYRQTITLDIP 360
Sbjct 3245840 TVETIDELNAAIKRADANDIIKFKPEKEKTINNSFSIETKKTVTIELDGRYRQTITLDIP
    
```

```

Sbjct 3245840 TVETIDELNAAIKRADANDIIKFKPEKEKTINNSFSIETKKTVTIELDGRYRQTITLDIP 3245661
Query 361      NGKFNNYAEIEGGVKLKNIKNESLVNKGSIQDLDIYDENGCKIENESSGEIWFVTIVEEA 420
Sbjct 3245660 NGKFNNYAEIEGGVKLKNIKNESLVNKGSIQDLDIYDENGCKIENESSGEIWFVTIVEEA 3245481
Query 421      NDVYIVNSGDITKISNNSSTIIRNSGNI DTVTGKKEPAISGNPKVNDTEKETKAARGL 480
                NDVYIVNSGDITKISNNSSTIIRNSGNI DTVTGKKEPAISGNPKVNDTEKETKAARGL
Sbjct 3245480 NDVYIVNSGDITKISNNSSTIIRNSGNI DTVTGKKEPAISGNPKVNDTEKETKAARGL 3245301
Query 481      NPRVEACSVPKDYVMITIPNSPKDSRYKIYYRVVYNKPYAMDVGDKINIGEWTVAPTDE 540
                NPRVEACSVPKDYVMITIPNSPKDSRYKIYYRVVYNKPYAMDVGDKINIGEWTVAPTDE
Sbjct 3245300 NPRVEACSVPKDYVMITIPNSPKDSRYKIYYRVVYNKPYAMDVGDKINIGEWTVAPTDE 3245121
Query 541      EPFLEKAKNGCYVEAVEVNTSTKEVSRWGRTNAT 574
                EPFLEKAKNGCYVEAVEVNTSTKEVSRWGRTNAT
Sbjct 3245120 EPFLEKAKNGCYVEAVEVNTSTKEVSRWGRTNAT 3245019

```

Figure 3. 15. Alignment of the predicted protein sequence of r-cwp66 with cwp66 630 (A) and BLAST output from a search of the protein database (B).

A

```

srB      MKKLYRIVINIILVLVILYSGFNIYSKLTQYNHDTKISSELQKKEYKKEDLSKINSDFKF 60
rsrB     -----NHDTKISSELQKKEYKKEDLSKINSDFKF 29
          *****

srB      WLSVENTNINYPVVQSKDNSYYLDKDFYKKDSISGTLFMDYRNKSIDDKNIIYGHNMKN 120
rsrB     WLSVENTNINYPVVQSKDNSYYLDKDFYKKDSISGTLFMDYRNKSIDDKNIIYGHNMKN 89
          *****

srB      KTMFNLNKFKDADFFKKNKIKITLNGKEFLYDVFSAYIVESDYDYLKTNFNNESDYQN 180
rsrB     KTMFNLNKFKDADFFKKNKIKITLNGKEFLYDVFSAYIVESDYDYLKTNFNNESDYQN 149
          *****

srB      YINDITSKSLYKSPIKVNNDKIVTLSTCTYEFDDARMVIHGRLI 225
rsrB     YINDITSKSLYKSPIKVNNDKIVTLSTCTYEFDDARMVIHGR-- 192
          *****

```

B  [ref|NC_009089.1](#)  Clostridium difficile 630 chromosome,
complete genome
Length=4290252

Sort alignments for this subject sequence by:
E value Score Percent identity Query start position Subject start position

Features in this part of subject sequence:
[putative sortase B](#)

Score = 323 bits (829), Expect = 1e-88, Method: Compositional matrix adjust.
Identities = 192/192 (100%), Positives = 192/192 (100%), Gaps = 0/192 (0%)
Frame = +3

Query	1	NHDTKISSELQKKEYKKEDLSKINSDFKFWLSVENTNINYPVVQSKDNSYYLDKDFYKKD	60
Sbjct	3148470	NHDTKISSELQKKEYKKEDLSKINSDFKFWLSVENTNINYPVVQSKDNSYYLDKDFYKKD	3148649
Query	61	SISGTLFMDYRNKSIDDKNIIYGHNMKNKTMfnlnkfkdadffkknkikitlngKEF	120
Sbjct	3148650	SISGTLFMDYRNKSIDDKNIIYGHNMKNKTMFNLNKFKDADFFKKNKIKITLNGKEF	3148829
Query	121	LYDVFSAYIVESDYDYLKTNFNNESDYQNYINDITSKSLYKSPIKVNNDKIVTLSTCTY	180
Sbjct	3148830	LYDVFSAYIVESDYDYLKTNFNNESDYQNYINDITSKSLYKSPIKVNNDKIVTLSTCTY	3149009
Query	181	EFDDARMVIHGR	192
Sbjct	3149010	EFDDARMVIHGR	3149045

Figure 3. 16. Alignment of the predicted protein sequence of r-sortaseB with sortaseB630 (A) and BLAST output from a search of the protein database (B).

3.2.3 Expression of the target proteins

One positive clone for each target was selected and plasmid DNA was purified from cultures of *E. coli* NovaBlue cells. Purified plasmid DNA was transformed into competent *E. coli* BL21 (DE3), the strain to be used for expression, and colonies from overnight agar cultures were picked for target expression studies. Each was tested in the presence or absence of IPTG inducer (Figure 3.17). Dot blots from the cultures were analysed by probing for the S.tag sequence. This revealed that the recombinant proteins could be expressed even without IPTG. Expression of Cwp84 in dot blots showed weak signal in the absence of IPTG and after the addition of the inducer. Whilst signal intensity for many other targets was good under non-induced and induced conditions, expression of the N terminal part of Cwp66 showed some dependence upon IPTG (column 3) as did expression of FbpA (column 4).

Specifically, GroEL was expressed even without IPTG and the signal from dot blots for both cultures was intense. Column 6 shows FliD expression which can be described as semi intense in signal under the two experimental conditions. In the next column, the expressed protein was Acd where IPTG did not appear necessary for protein production. In the next three columns, (FliC, C-terminus of Cwp66 and Sortase B) very good signal strength was observed. The last column shows that expression of the full-length Cwp66 in the presence IPTG was stronger than in the absence of inducer (Figure 3.16).

Using this information, large cultures of BL21 (DE3) cells were prepared for expression of each target, analysis in Western blots and purification of the target proteins. Before purification, a Western blot was run for all extracted samples from the bacterial cells. The data are shown in Figures 3.18 and 3.19.

In Figure 3.18, lane 1 shows the expression of Cwp84 but no band indicating the presence of the recombinant protein can be seen. In contrast, lane 2 shows a strong band at a molecular weight over 22 kDa corresponding to CspA (Table 3.2; predicted molecular weight 27.5 kDa). In the same Figure, a band just under 50 kDa in lane 3 shows expression of the N-terminus of Cwp66, a recombinant protein with a predicted molecular weight of 47.1 kDa (Table 3.2). Lane 4 shows that like lane 1, no signal was detected to show expression of a protein corresponding to FbpA. In lanes 5 and 6 of Figure 3.18, two bands in excess of 64 kDa show the expression of GroEL and Flid (72.2 and 75.8 kDa; Tables 3.2).

Figure 3.19 shows the analysis of bacterial extracts for evidence of the expression of other target proteins. In lane 1, there was no evidence for the successful expression of Acd but a band of good signal strength can be seen in lane 2 (Flc, 48 kDa; Table 3.2). Two strong bands are also visible in lanes 3 and 4 showing the successful expression of the C-terminus of Cwp66 and Sortase B respectively (predicted molecular weights of 53.9 and 43.4 kDa, Table 3.2). The last lane of Figure 3.19 shows a band over 64 kDa likely to be Cwp66.

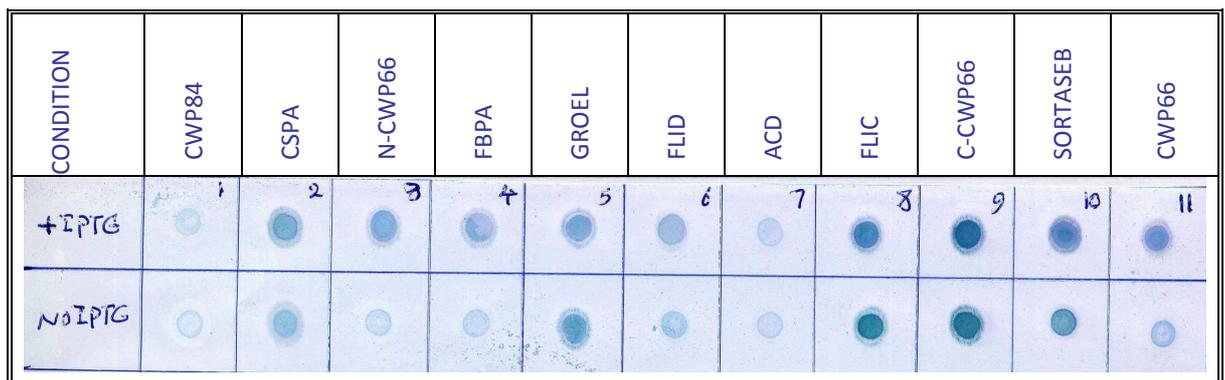


Figure 3.17. Dot blot analysis of the expression of target proteins in *E. coli* BL21(DE3) in the presence or absence of IPTG.

After application of bacterial extracts to the membrane and blocking, recombinant proteins were detected with anti S-tag antibody, anti-mouse-HRP and TMB substrate.

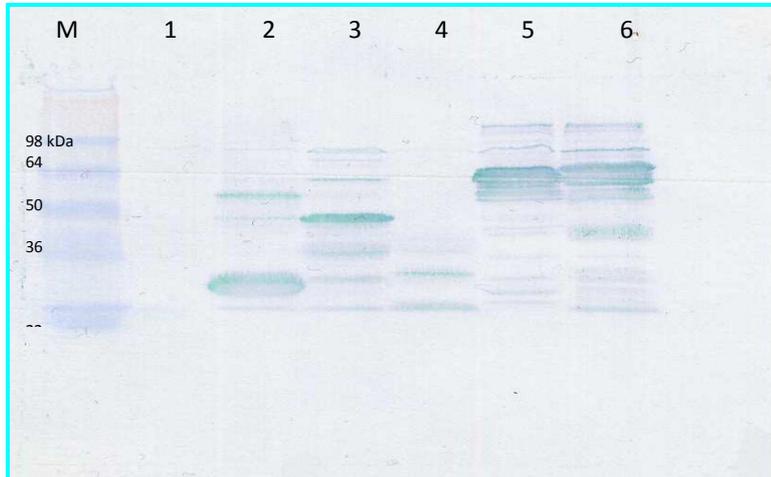


Figure 3. 18. Western blot analysis of extracts of *E. coli* BL21(DE3) cells expressing six target proteins

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, Cwp84; lane 2, CspA; lane 3, N-terminus of Cwp66; lane 4, FbpA; lane 5, GroEL; lane 6, FliD.

After transfer and blocking, recombinant proteins were detected with anti S-tag antibody, anti-mouse-HRP and TMB substrate.

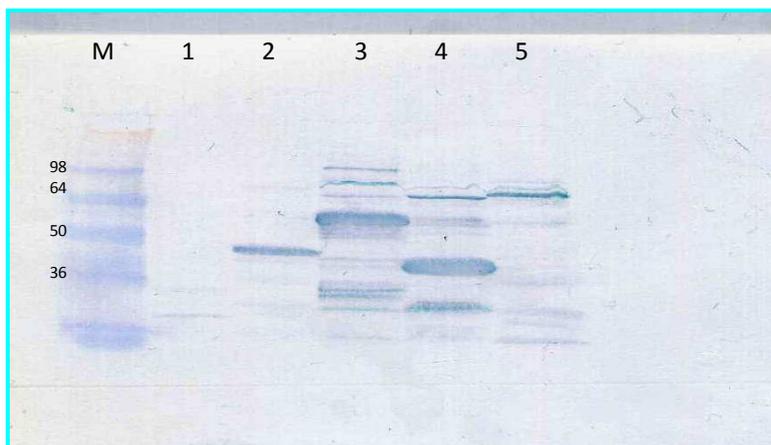


Figure 3. 19. Western blot analysis of extracts of *E. coli* BL21(DE3) cells expressing five target proteins

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, Acd; lane 2, FliC; lane 3, C-terminus of Cwp66; lane 4, SortaseB; lane 5, Cwp66.

After transfer and blocking, recombinant proteins were detected with anti S-tag antibody, anti mouse-HRP and TMB substrate.

Inserts of 6 sequences – *cwp84*, *fbpA*, *acd*, *fliC*, 3' *cwp66* and *cwp66* – were not expressed particularly well from the pET32 vector system, as judged by Western blotting. In an attempt to overcome this problem, plasmid DNA was transformed into *E. coli* BL21(DE3) pLysS and C43 (DE3) competent cells. Results from both the new expression hosts is shown in Figure 3.20 and 3.21. In Figure 3.20, lanes 1 and 2 are samples from bacterial expressing Cwp84 and FbpA; the blots do not show signals that could be attributed to the recombinant proteins. Lane 3 in Figure 3.20 shows a band over 64 kDa that could be the recombinant Acd. In contrast, there is clear evidence of a band in lane 4 of same Figure close to 50 kDa that might represent FliC (predicted molecular weight 48 kDa, Table 3.2). The results in lane 5 are more complex but amongst the many bands, a particularly strong signal can be seen a molecular weight over 50 kDa. As with FliC expression, single, clear band appears in lane 6 of Figure 3.20 that might represent Cwp66 (predicted molecular mass 83.3 kDa).

Further attempts were made to express these more problematic target proteins by transforming the relevant constructs into *E. coli* C43 (DE3). The results of expression experiments and Western blotting analysis are shown in Figure 3.21. Once again, there was no evidence of successful expression of Cwp84 and FbpA, (lanes 1, 2 and 5) and a clear band over 64 kDa (lane 3) suggested that recombinant Acd had been produced. Lanes 4 and 6 illustrate the production of the C-terminus of Cwp66 with good signal strength from a protein over 50 kDa in molecular weight. In Figure 3.21 two separate transformants were prepared for constructs of FbpA (lanes 2 and 5) and C-terminus of Cwp66 (lanes 4 and 6).

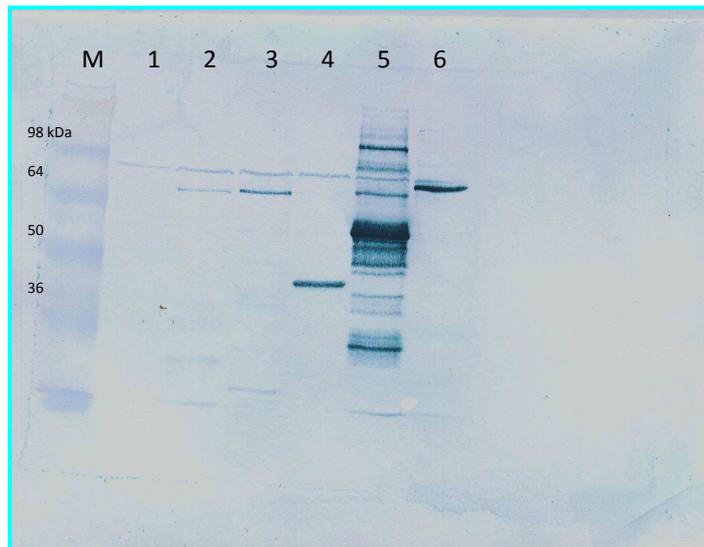


Figure 3. 20. Western blot analysis of extracts of *E. coli* BL21(DE3) pLysS expressing clostridial target proteins

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, Cwp84; lane 2, FbpA; lane 3, Acd; lane 4, FliC; lane 5, C-terminus of Cwp66; lane 6, Cwp66

After transfer and blocking, recombinant proteins were detected with anti S-tag antibody, anti mouse-HRP and TMB substrate.

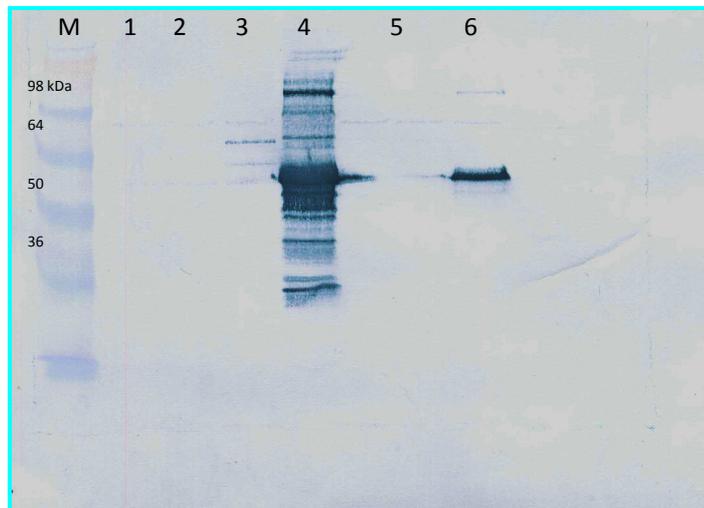


Figure 3.21. Western blot analysis of extracts of *E. coli* C43(DE3) expressing clostridial target proteins

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, Cwp84; lane 2, FbpA(1st transformant); lane 3, Acd; lanes 4 and 6(two transformants of target), C-terminus of Cwp66; lane 5, FbpA(2nd transformants)

3.2.4 Purification of recombinant proteins

Recombinant proteins produced from the pET32 expression system in *E. coli* strains BL21(DE3), BL21(DE3) pLysS and C43(DE3)] were purified by Ni-chelation chromatography. For each protein, lysates were produced from cultures induced overnight with IPTG and these were passed through a 5 ml pre-packed column that had been equilibrated previously with binding buffer. The columns were washed with binding buffer and attached proteins were eluted with 200-300 mM imidazole in 10 fractions. Fractions from each column were analysed by SDS-PAGE.

3.2.4.1 Purification of Cwp84

Cwp84 was transformed into all three *E. coli* strains BL21(DE3), BL21(DE3) pLysS and C43(DE3) but a recombinant protein of the expected protein molecular weight (102 kDa) could not be detected under any of the experimental conditions tested.

3.2.4.2 Purification of recombinant CspA

In order to obtain sufficient recombinant CspA for further experimentation, different culture volumes were tested. Lysates from 100 and 200 ml cultures yielded better results than bigger culture volumes. For cultures of these sizes, lysates were passed through a column that had been first equilibrated with binding buffer. The column was washed with 50 ml of this buffer and further treated with 50 ml of binding containing 20 mM imidazole to remove proteins that were loosely associated. Proteins were then recovered with higher concentrations of imidazole. Figure 3.22 shows the purification of CspA from lysate of a 100 ml culture. The flow through after loading (lane 1) contained a significant number of proteins and

a range of species were eliminated from the column by washing (lanes 2 and 3). Elevation of the imidazole concentration to 200 mM initially recovered a protein of the predicted molecular weight along with contaminating species (lane 4). Subsequent fractions contained a high concentration of the recombinant protein with modest contamination (lanes 5 and 6) before concentration declined (lanes 7 to 9). The molecular weight of the purified protein (27 kDa) was consistent with that expected for CspA with the purification and detection tags encoded by pET32 (Table 3.2).

In an effort to enhance the purity of the protein and its concentration, the method was repeated with extra washing using buffer containing imidazole. Figure 3.23 shows that extensive washing with imidazole rising to 100 mM essentially eliminated contaminating proteins; CspA was then eluted in high concentration with 200 mM imidazole. In Figure 3.23, the protein-charged column was washed with binding buffer containing 40 mM imidazole (lane 4), 60 mM (lanes 5 to 10), and 100 mM (lanes 11 to 16). CspA was eluted with a sharp profile on raising the imidazole to 200 mM (lanes 17 to 21). Those fractions that appeared to contain a single band for CspA (Figure 3.22 lanes 5 to 9; Figure 3.23 lanes 18 to 21) were pooled and dialysed in preparation for later experiments.

3.2.4.3 Purification of recombinant N-terminus of Cwp66

Based on the primer design and tags present in pET32, the expressed protein was expected to have a molecular weight of 47.1 kDa (Table 3.2). To purify this protein, lysates prepared from induced cultures of different volumes were applied to an equilibrated column. Figure 3.23 shows that a wide range of proteins were washed from the column with binding buffer (lane 1) and 40 mM imidazole (lane 2). The recombinant protein was eluted on switching to 200 mM imidazole. Initial recoveries were contaminated with other proteins (lane 4) but purity improved in later fractions (lane 5). An intense band with a molecular weight close to 50 kDa

was consistent with that predicted for the purified N-terminal region of Cwp66. The purity of the sample in lane 5 was judged sufficient for the purposes of later experiments.

3.2.4.4 Purification of FbpA

None of the *E. coli* strains tested expressed a protein likely to be FbpA (predicted molecular weight of 85.2 kDa; Table 3.2).

3.2.4.5 Purification of GroEL

As in other purifications, lysates from induced cultures expressing GroEL were applied to columns and eluted with different concentration of imidazole in binding buffer. The impact of different concentrations of imidazole were explored but the best results were obtained using 200 mM imidazole in binding buffer. Figure 3.25 shows the presence of a prominent protein over 64 kDa in molecular weight, the fraction loaded to lane 6 showing better purity than that eluted earlier (lane 5). The expected molecular weight for this protein is 72.2 kDa (Table 3.2). This is consistent with the prominent protein visible in lane 6 of Figure 3.25.

3.2.4.6 Purification of FliD:

The expected band for FliD was 75.8 kDa (Table 3.2) and a protein of this size was obtained by purification on Ni-chromatography. Two lysates were run on columns to purify the recombinant FliD using different concentrations of imidazole in the binding buffer. Figure 3.26 shows a stepwise elution of FliD with different concentration of imidazole from 60 to 300 mM over 30 fractions. This Figure reveals that although elution of FliD started at 60 mM imidazole, 200 mM of imidazole was the best concentration with intense bands evident in SDS-PAGE. Lane 1 shows that a great many proteins were present in the original lysate but amongst these is a clear band that could be FliD (75.8 kDa). The column flow through is shown in lane 2 and washing with 40 mM imidazole (lane 3) shows no

evidence of release of a protein similar in size to that expected for FliD. A shift to 60 mM imidazole in binding buffer initially showed no sign of release of FliD (lanes 4 and 5) but a protein that could be the recombinant proteins appears in later fractions (lanes 6-9). The next 7 fractions (lanes 10-16) were eluted with 100 mM imidazole in binding buffer. Although some FliD may be lost from the column, contaminants of lower molecular weight are also eliminated. Increasing the concentration of imidazole to 200 mM shows the recovery of a protein of the predicted molecular weight of FliD (lanes 17 to 27). The recombinant protein is reasonably pure, although slight contamination with a species of around 36 kDa can be seen in some lanes. The final elution step used buffer containing 300 mM imidazole. FliD was recovered (lanes 28 to 33) but the concentration was low.

Based on this experiment, a second column was loaded with the bacterial lysate, washed with binding buffer containing 40 mM imidazole and eluted with 200 mM imidazole. Figure 3.27 shows the success of this protocol; FliD was isolated in reasonable purity. Lanes 3 to 9 show the fractions eluted with 200 mM imidazole in binding buffer. Immediately after the change of buffer, no proteins were released (lane 3). Following this, proteins appeared that were slightly lower in molecular weight than FliD (lane 4), resolving to FliD and a smaller protein (lane 5). Thereafter, the dominant species was FliD (lanes 6-9).

3.2.4.7 Purification of Acd:

The lysates of bacterial cultures expressing Acd were applied into the column in different volumes and under different conditions but it proved impossible to purify the recombinant protein. This was consistent with the very low signal obtained in dot blots (Figure 3.17).

3.2.4.8 Purification of FliC:

Recombinant FliC was purified in the same way as other recombinant proteins; the loaded column was washed first with binding buffer, and then 40 mM imidazole before a solution of binding buffer with 200 mM imidazole was used to elute the attached protein. Figures 3.28 and 3.29 show a clear band just less than 50 kDa which is the expected size for FliC (48.0 kDa, Table 3.2). Both Figures show that FliC was not eluted in a completely pure state. In Figure 3.28, lanes 1 and 2 show samples for the bacterial lysate and the column flow through. Washing the column with binding buffer failed to release protein (lane 3) while lane 4 shows a range of proteins were washed out with 40 mM imidazole. Other lanes in Figure 3.28 illustrate elution of attached proteins with 200 mM imidazole.

Although a prominent protein of the right molecular weight can be seen in lane 7, there are many other contaminants present. The later fraction (lane 8) shows a relatively pure band close to 50 kDa but the concentration is low. Another column was set up to attempt to purify FliC using the same conditions (Figure 3.29). SDS-PAGE analysis of all samples from the second attempt showed only one fraction (Figure 3.29, lane 18) with a band likely to correspond to the recombinant protein.

3.2.4.9 Purification of C-terminus of Cwp66

In purifying the C-terminal region of Cwp66, a protein of 53.9 kDa was sought (Table 3.2). The loaded column was first washed with binding buffer and then 40 mM imidazole (each 50 ml) to remove loosely-associated proteins (samples not shown). Buffer containing 200 mM imidazole was then used for elution. Figure 3.30, lanes 1 to 8 shows analysis of fractions eluted in this way. Although the prominent band evident in these samples was higher in molecular weight than expected, samples shown in lanes 4 to 7 were pooled and dialysed in preparation for the next phases of the analysis, given that these fractions contained fewer contaminating proteins than others (*eg* the species evident in lane 8).

3.2.4.10 Purification of Sortase B

Sortase B is another protein which was expressed and purified by Ni chromatography. Table 3.2 shows the expected size for Sortase B should be 43.4 kDa. SDS-PAGE analysis of samples from purification of recombinant Sortase B can be seen in Figure 3.31. The equilibrated column was loaded with bacterial lysate (lane 1) and washed with binding buffer (50 ml; lane 2). The original lysate shows a strong sharp band under 50 kDa that might represent Sortase B. Washing with binding buffer (lane 2) removed a range of unbound proteins of variable molecular weight and this was enhanced by washing with 40 mM imidazole (lane 3). Preliminary experiments with eluting concentrations of imidazole between 100 and 300 mM showed (data not shown) 200 mM imidazole to be the optimum concentration to elute Sortase B. Therefore this solution was used (lanes 4 to 9) to recover 6 fractions containing a protein of the expected mass. These fractions were mixed and dialysed for the next step.

3.2.4.11 Purification of Cwp66

The last recombinant protein to be expressed and purified was Cwp66 (predicted mass of 83.3 kDa, Table 3.2). Purification of this recombinant protein was carried out using very similar methods to those already described. The equilibrated column was firstly was loaded with bacterial lysate and then washed with binding buffer. Lane 1 in Figure 3.32 shows a strong band of the expected molecular weight for recombinant Cwp66 was present in the original lysate and that this was released from the column by washing with binding buffer (lane 2) and 40 mM imidazole (lane 3). After washing, the column was eluted with 200 mM imidazole to collect fractions. Lanes 4 to 9 in Figure 3.32 show the analysis of fractions 9-14 from this elution. A prominent band of a size consistent with Cwp66 can be seen. These fractions were pooled and dialysed for the next step.

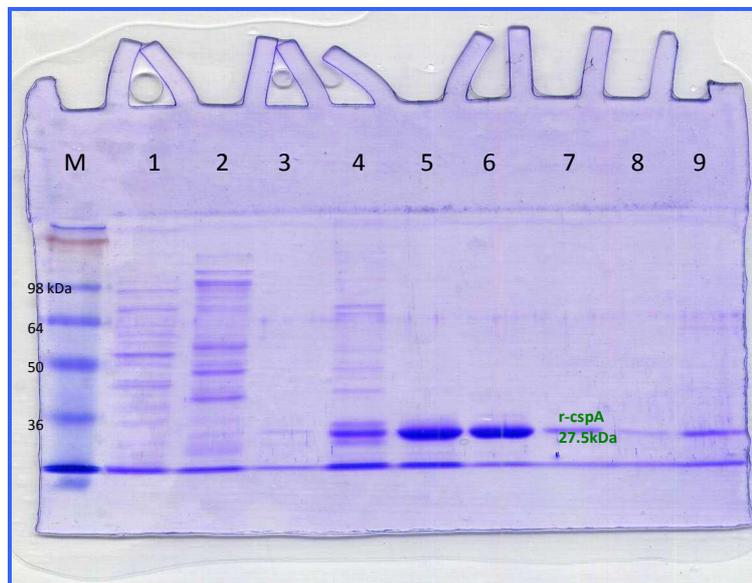


Figure 3. 22. SDS-PAGE analysis of the purification of recombinant CspA

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1; flow through from loading of lysate to the column; lane 2; washings recovered with binding buffer; lane 3; washings recovered with 40mM imidazole in binding buffer; lanes 4 to 9, fractions eluted with 200 mM imidazole in binding buffer.

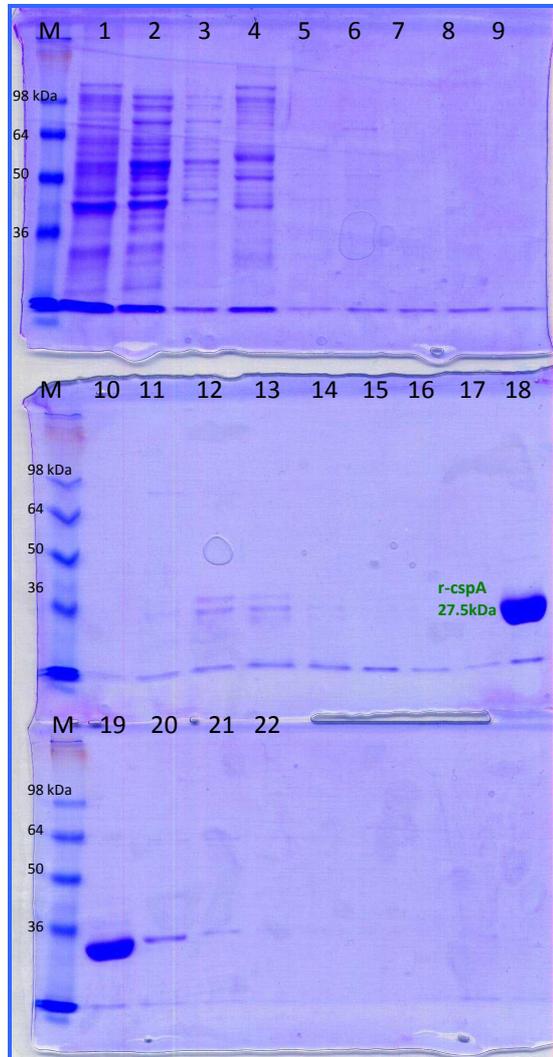


Figure 3. 23. SDS-PAGE analysis of the purification of recombinant CspA

Upper gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, original lysate; lane 2, flow through from the column on loading; lane 3, washings recovered with binding buffer; lane 4, washings recovered with 40 mM imidazole; lanes 5 to 9, washings recovered with 60 mM imidazole.

Middle gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 10, washings recovered with 60 mM imidazole; lanes 11 to 16, washings recovered with 100 mM imidazole; lanes 17 to 18, fractions eluted with 200 mM imidazole

Lower gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lanes 19 to 21, fractions eluted with 200 mM imidazole.

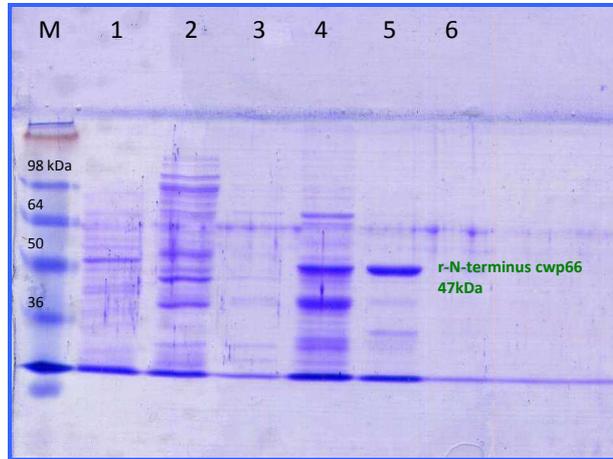


Figure 3. 24. SDS-PAGE analysis of the purification of recombinant N-terminus Cwp66

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1; washings recovered with binding buffer; lane 2, washings recovered with 40 mM imidazole in binding buffer; lanes 3 to 6, fractions eluted with 200 mM imidazole in binding buffer.

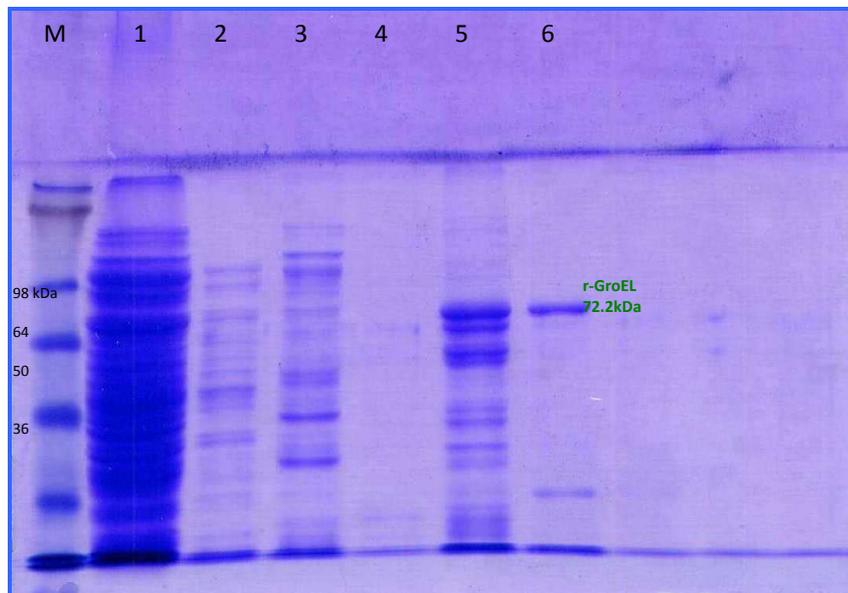


Figure 3. 25. SDS-PAGE analysis of the purification of recombinant GroEL

From the left: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, original lysate; lane 2; washings recovered with binding buffer lane 3, washings recovered with 40 mM imidazole in binding buffer; lanes 4 to 6; fractions eluted with 200 mM imidazole in binding buffer.

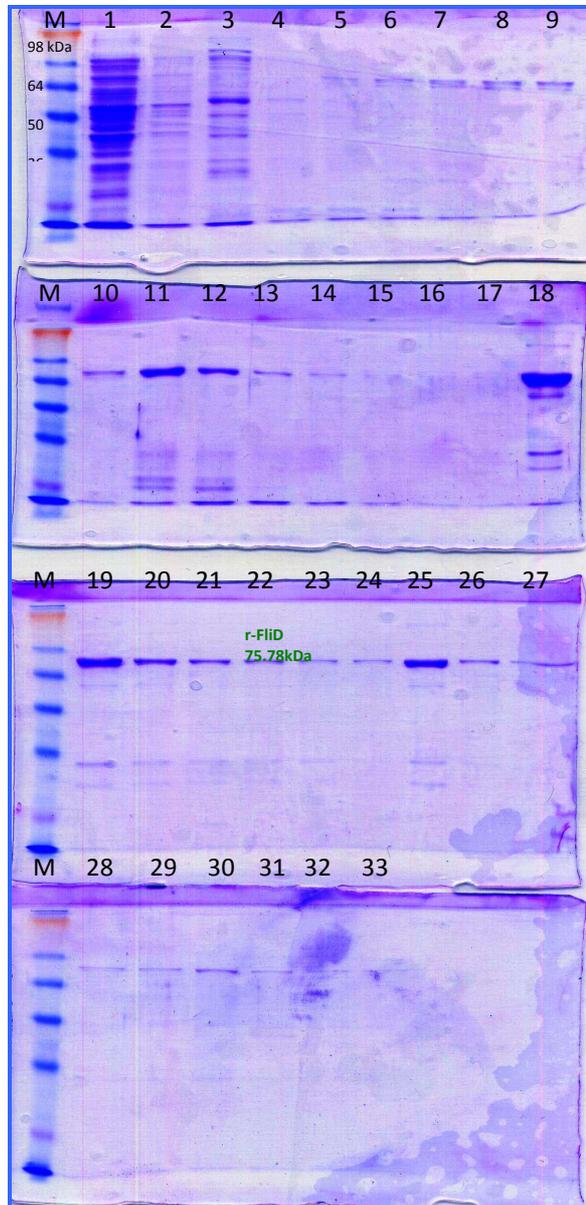


Figure 3. 26. SDS-PAGE analysis of the purification of recombinant FliD

Top gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lane 1, sample before application to the column; lane 2, flow through from the column after loading; lane 3, washings recovered with 40 mM imidazole; lanes 4 to 9, washings recovered with 60 mM imidazole

Second gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lanes 10 to 16, washings recovered with 100 mM imidazole, lanes 17 to 18, fractions eluted with 200 mM imidazole

Third gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lanes 19 to 27, fractions eluted with 200 mM imidazole.

Bottom gel: M, SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen); lanes 28 to -33, fractions elution with 300 mM imidazole.

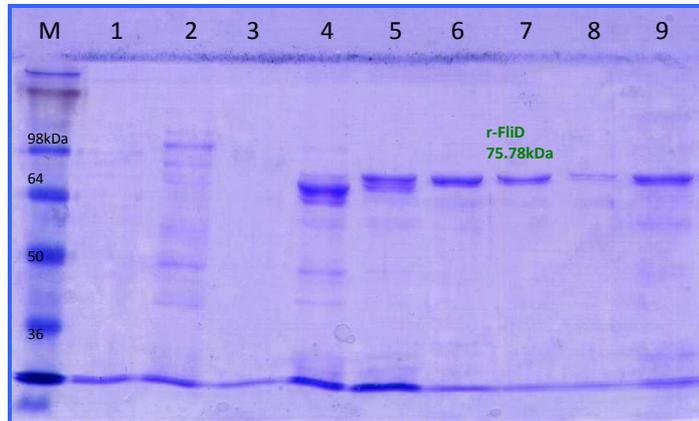


Figure 3. 27. SDS-PAGE analysis of purification of recombinant FliD

From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lane 1, flow through from the column after loading; lane 2, washings recovered with 40 mM imidazole ; lanes 3 to 9, fractions eluted with 200 mM imidazole.

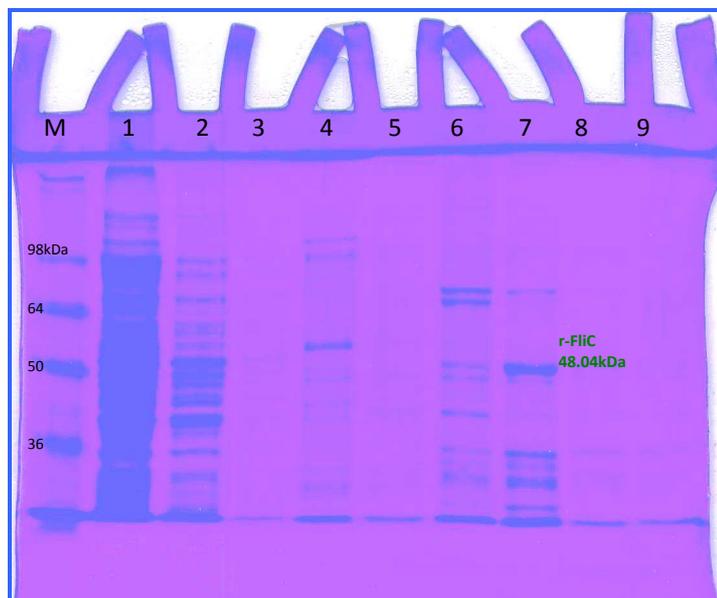


Figure 3. 28. SDS-PAGE analysis of purification of recombinant FliC

From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lane 1, original lysate; lane 2, flow through from the column after loading ; lane 3, washings recovered with binding buffer; lane 4, washings recovered with 40mM imidazole; lanes 5 to 9, fractions eluted with 200 mM imidazole.

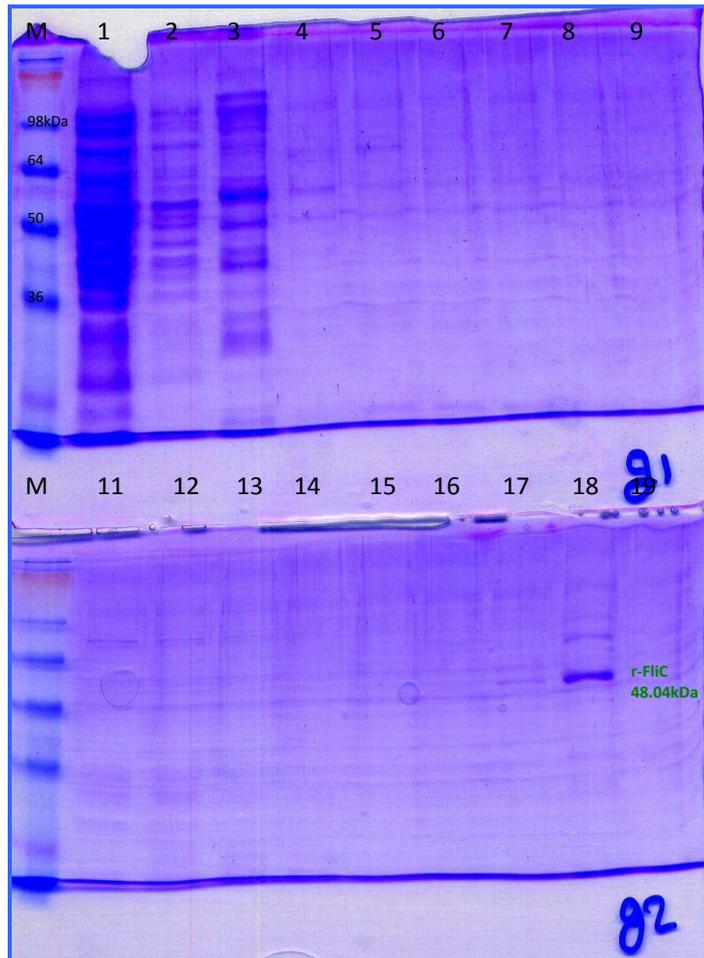


Figure 3. 29. SDS-PAGE analysis of purification of recombinant FliC

Gel 1: From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lane 1, original lysate; lane 2, washing recovered with binding buffer; lane 3, washing recovered with 40mM imidazole; lanes 4 to 9, fractions eluted with 200 mM imidazole.

Gel2; From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lanes 11 to 19, fractions recovered with 200 mM imidazole.

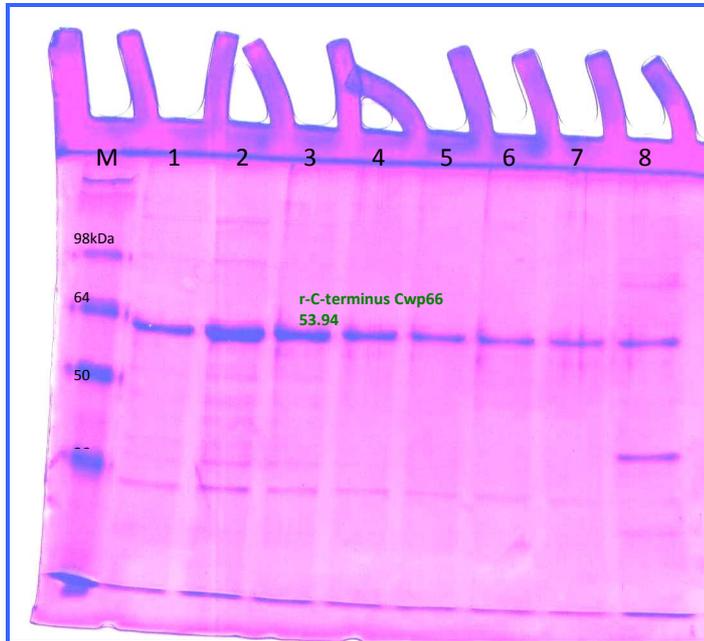


Figure 3. 30. SDS-PAGE analysis of purification of recombinant C-terminus Cwp66

From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lanes 1 to 8, fractions eluted with 200 mM imidazole (fractions 15-22).

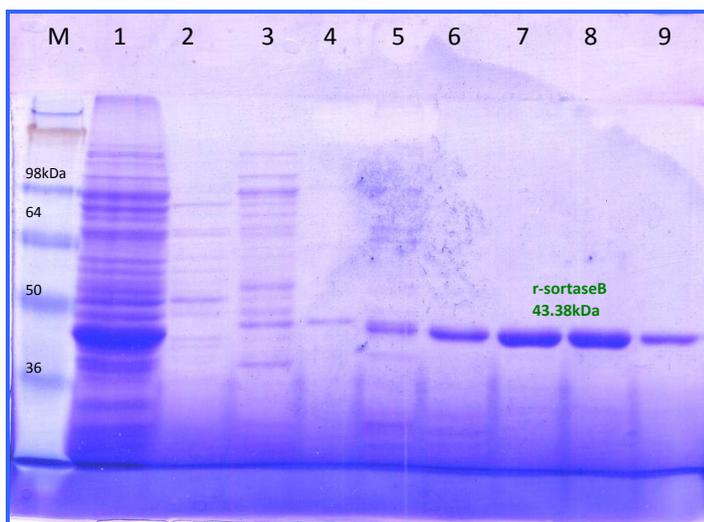


Figure 3. 31. SDS-PAGE analysis of purification of recombinant Sortase B

From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lane 1, original lysate; lane 2, washing recovered with binding buffer; lane 3, washing recovered with 40 mM imidazole; lanes 5 to 9, eluted fractions with 200 mM imidazole (fractions 1-6).

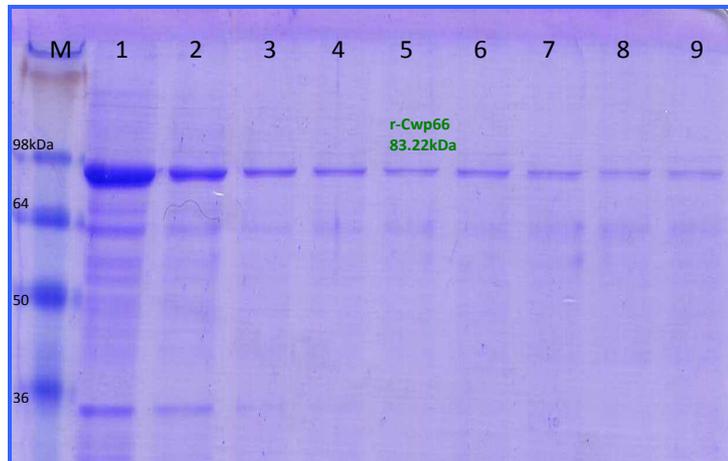


Figure 3. 32. SDS-PAGE analysis of purification of recombinant Cwp66

From left: M, SeeBlue Plus2 Pre-Stained molecular weight markers (Invitrogen); lane 1, original lysate; lane 2, washing recovered with binding buffer; lane 3, washing recovered with 40 mM imidazole; lane 4-9, eluted fractions with 200 mM imidazole (fractions 9 to 14).

3.2.5 Mass spectrometry of recombinant proteins

To confirm the identities of the expressed proteins and their similarity with the intended target proteins from *C. difficile* 630, a sample of each of the purified proteins was analysed by mass spectrometry. The single band related to each target was cut from an SDS-PAGE gel and sent to a central proteomic facility at the University of Glasgow where an LC-MS/MS system was used to analyse the samples. Mascot search results showed that most matched successfully with proteins from *C. difficile* 630. However the sequence coverage was variable for different recombinant proteins which could be explained by their lysine content. Hence, in this phase of the project, all the targets which were successfully expressed and purified were confirmed as the intended proteins and were related to their counterparts from *C. difficile* 630 with the exception of those proteins (Cwp84, FbpA and Acd) that could not be expressed in an *E. coli* host. The following Figures show the results from mass spectrometry analysis of the expressed targets (Figures 3.33-3.40). In all Figures, the matched peptides are in red and bold text.

Match to: [gi|126698469](#) Score: **120**
cold shock protein [Clostridium difficile 630]
 Found in search of F:\Naziri-7-12-9.wiff

Nominal mass (M_r): **7448**; Calculated pI value: **4.75**
 NCBI BLAST search of [gi|126698469](#) against nr
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Clostridium difficile 630](#)
[gi|115249906](#) from [Clostridium difficile 630](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: **25%**

Matched peptides shown in **Bold Red**

1 MKNQIVKWFN NEKGFGFISV EGEDDVVVFH SAIQNDGYK**T LEEGKVSFD**
 51 **ITQGNR**GPQA ENVNRI

Figure 3. 33. Result of Mascot search using data from mass spectrometry analysis of recombinant CspA.

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 25% of the sequence of native CspA could be recovered from the recombinant protein.

Match to: [gi|11066029](#) Score: **786**
Cwp66 [Clostridium difficile]
 Found in search of F:\Naziri-7-12-9.wiff
 Nominal mass (M_r): **66895**; Calculated pI value: **5.19**
 NCBI BLAST search of [gi|11066029](#) against nr
 Unformatted [sequence string](#) for pasting into other applications
 Taxonomy: [Clostridium difficile](#)
 Fixed modifications: Carbamidomethyl (C) Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: **20%**

Matched peptides shown in **Bold Red**

1 MNISKQIVSL LTMTFLT VTL YGNTSNASTK DTLTGSGRWE TAIKISQAGW
 51 **TKSEAVLVN DNSIADALSA TPF**AKAKDAP **ILLTQSNKLD** SRTKAEKRL
 101 GVKNVYLIGG SIALSSEIEK **QLNAENINFE** RISGNSRYDT SLKLAEKLDR
 151 EKSISKIVVV NGEK**GLADAV SVGAIAAQEN MPIILSDSEN GTEVADNFID**
 201 **SKDIAKSYVI GGTYSISNSV ER**SLPNATRI AGSSRSETNA **KIIEEFYKDT**
 251 **DIK**NIYVTKD **GTKNKNDLID SLAVGVLAAK** NSSPIVLGN KLDTTQKDVL
 301 NTKIIDKVTQ IGGLGNENVV EDILDIQEET KYTVETIDEL NAAIKRADAN
 351 DIIKFKPEKE KTINNSFSIE TKKTVTIELD GRYRQTITLD IPNGKFNNYA
 401 EIEGGVKLKN IKNESLVNKG SIQDLDIYDE NGCKIENESS GEIWFVTIVE
 451 EANDVYIVNS GDITKISNNS SSTIIRNSGN IDTVTGGKKEP AISGNKPKVN
 501 DTEKETKAAR GLNPRVEACS VPKKDYVMIT IPNSPKDSRY KIYYRVVYNK
 551 PYAMDVGDKI NIGEWTVAPT DEEPFLEKAK NGCYVEAVEV NTSTKEVSRW
 601 GRNATDDGF

Figure 3. 34. Result of Mascot search using data from mass spectrometry analysis of recombinant N-terminus of Cwp66.

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 20% of the sequence of native N-terminus of Cwp66 could be recovered from the recombinant protein.

Match to: gi|126697767 Score: 1583
60 kDa chaperonin [Clostridium difficile 630]
 Found in search of F:\Naziri-7-12-9.wiff
 Nominal mass (M_r): 57735; Calculated pI value: 4.74
 NCBI BLAST search of gi|126697767 against nr
 Unformatted [sequence string](#) for pasting into other applications
 Taxonomy: [Clostridium difficile 630](#)
 Links to retrieve other entries containing this sequence from NCBI Entrez:
[gi|118597108](#) from [Clostridium difficile 630](#)
[gi|115249204](#) from [Clostridium difficile 630](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: 45%
 Matched peptides shown in **Bold Red**

1 MAKEIKFSEE TRRALEAGVN KLADTVKVTL GPKGRNVILD KCFGSPPLITN
 51 DGVTIAK**EIE LEDRFENMGA QLVKEVATKT NDVAGDGT TT ATVLAQAIIR**
 101 EGLKNVTAGA NPILLRKG IQ **KAVTVAVEEL KNQSRIVETQ E AISQV ASIS**
 151 **AGDEEVGKLI AEAMEIVGKD** GVITVEESQT MNT ELDAVEG MQFDR**GFVSA**
 201 **YMVTDVDKME AVLNDPYILI TDKKISNIQE LLPVLEQIVQ QGK KLLIAE**
 251 **DVEGEALSTL VVNKL RGTFD VVA VK**APGFG DRRKEM LQDI AILTGAQV IS
 301 EELGYDLKEA **DLSMLGRASS** VKVTKESTI **VDGSGDKKAI** ED RVTQIKHQ
 351 **VEQTTSDFDR EKL MERLAKL AGGVA VVKVG AATEVELKER** KLR**IEDALNA**
 401 **TRAAVEEGIV** AGGTA FVSV IPAIGTLIES LEGEVKLGAQ IVK KALEEPL
 451 RQIAINAGLE GAVIVQNVVN SEAETGFDAL NEK**YVNMIEA GIVDPTK**VSR
 501 SALQNAASIA STFLTTEAAV ADLPEKEDAG MPGMGGGMPG MM

Figure 3. 35. Result of Mascot search using data from mass spectrometry analysis of recombinant GroEL

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 45% of the sequence of native GroEL could be recovered from the recombinant protein.

Match to: gi|126697808 Score: 1515
flagellar cap protein [Clostridium difficile 630]
 Found in search of F:\Naziri-7-12-9.wiff
 Nominal mass (M_r): 56099; Calculated pI value: 5.37
 NCBI BLAST search of gi|126697808 against nr
 Unformatted [sequence string](#) for pasting into other applications
 Taxonomy: [Clostridium difficile 630](#)
 Links to retrieve other entries containing this sequence from NCBI Entrez:
[gi|13324551](#) from [Clostridium difficile](#)
[gi|115249245](#) from [Clostridium difficile 630](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: 60%
 Matched peptides shown in **Bold Red**

1 MSSISPVR**VT GLSGNFDMEG IEAS MIRDK** EKVNKAKQDQ QIVKWKQEIY
 51 RDIKESKNL YDKYLN**GDSP NSITNKKAYS ATRITSSDES IIVAKGSAGA**
 101 EK**INYQFAVS QMAEPAKVTI** KLNSSDPVQ QFPPN**ASGAS SLNIGGVNIP**
 151 **ISEQDTTSTI VSKINSLCAD NDIRASYSEM** TGELIISRKQ TGSSSDIDLK
 201 **VIGNDSL**AGQ IASDN**GITFT TDASG**TKSAV VYGKNLEADV TDDQGRVTHI
 251 SKEQNSFKID NIDYVNSKG SAKLVSV**TD T EATKNMKA F VDDYNALMDK**
 301 **VYGLVTTKKS** KDYPPLTDEQ KDDMT**TEEIE KWEKKAKEGI** LRNDDELRAF
 351 VEDIQSMFFG DADTIALR**K LGISEHENYN KKGQISFNAD TFSKALIDDS**
 401 **DKVYKALAGY** SSNYDDKGMF EKLK**KIVFEY SGSSASKLTK KAGMENS**SSA
 451 **SQNVYSKQIA** EQERNISRLV EKMNDKEKRL YAK**YSALES LNKYSSQ**MNY
 501 FSQAQGN

Figure 3. 36. Result of Mascot search using data from mass spectrometry analysis of recombinant FlID

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 60% of the sequence of native FlID could be recovered from the recombinant protein.

Match to: [gi|126697810](#) Score: **961**
flagellin subunit [Clostridium difficile 630]
 Found in search of F:\Naziri-7-12-9.wiff
 Nominal mass (M_r): **30755**; Calculated pI value: **7.88**
 NCBI BLAST search of [gi|126697810](#) against nr
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Clostridium difficile 630](#)
 Links to retrieve other entries containing this sequence from NCBI Entrez:
[gi|6456511](#) from [Clostridium difficile](#)
[gi|10281483](#) from [Clostridium difficile](#)
[gi|115249247](#) from [Clostridium difficile 630](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: **51%**

Matched peptides shown in **Bold Red**

1 MRVNTNVSAL IANNQMGRNV NGQSKSMEKL SSGVRIKRAA **DDAAGLAISE**
 51 **KMRAQIKGLD QAGRNVDQGI SVVQTAEGSL EETGNILQRM RTLSLQSAN**E
 101 **INNTEEREKI ADELTLQKDE IERISSSTEF NGKKLLDGTS STIRLQVGAS**
 151 **YGTNVSQTSN NNNEIKIQLV NTASIMASAG ITTASIGSMK AGGTTGTDA**A
 201 **KTMVSSLDAA LKSLNSSRAK LGAQQNRLES TQNNLNNTLE NVTAAESRIR**
 251 **DTDVASEMVN LSKMNLVQA SQSMLAQANQ QPQGVLLG**

Figure 3. 37. Result of Mascot search using data from mass spectrometry analysis of recombinant FliC

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 51% of the sequence of native FliC could be recovered from the recombinant protein.

Match to: [gi|11066029](#) Score: **826**
Cwp66 [Clostridium difficile]
 Found in search of C:\Documents and Settings\rjb12n\My Documents\Ali-14-5-10.wiff
 Nominal mass (M_r): **66895**; Calculated pI value: **5.19**
 NCBI BLAST search of [gi|11066029](#) against nr
 Unformatted [sequence string](#) for pasting into other applications
 Taxonomy: [Clostridium difficile](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: **23%**

Matched peptides shown in **Bold Red**

1 MNISKQIVSL LTMTFLTVTL YGNTSNASTK DTLTGSGRWE TAIKISQAGW
 51 TKSESAVLVN DNSIADALSA TPFKAKDAP ILLTQSNKLD SRTKAEKRL
 101 GVK**NVYLIGG SIALSSEIEK** QLNAENINFE RISGNSRYDT SLKLAEKLDLDR
 151 EKSISKIVVV NGEKGLADAV SVGAIAAQEN MPIILSDSEN GTEVADNFID
 201 SKDIAKSYVI GGTYSISNSV ERSLPNATRI AGSSRSETNA KIIEEFYKDT
 251 DIKNIYVTKD GTKNKNLID SLAVGVLA AK NSSPIVLGN KLDTTQKDV L
 301 NTKIIDKVTQ IGGLGNENVV EDILDIQEET **KYTVETIDEL NAAIKRADAN**
 351 DIIKFKPEKE **KTINNSFSIE TKKTVFIELD GR YRQTITLD IPNGKFNNYA**
 401 **IEGGVKLKN** IKNESLVNKG SIQDLDIYDE NGCKIENESS GEIWFVTIVE
 451 EANDVYIVNS GDITK**ISNNS SSTIIRNSGN IDTVTGKKEP** AISGNKPKVN
 501 DTEKETKAAR GLNPRVEACS VPK**KDYVMIT IPNSPKDSRY** KIYYR**VVYNK**
 551 **PYAMDVGDKI NIGEWTVAPT DEEPFLEK**AK NGCYVEAVEV NTSTKEVSRW
 601 GRNATDDGF

Figure 3. 38. Result of Mascot search using data from mass spectrometry analysis of recombinant C-terminus of Cwp66.

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 23% of the sequence of native Cwp66 could be recovered from the recombinant protein.

Match to: [gi|11066029](#) Score: 1384
Cwp66 [Clostridium difficile]
 Found in search of F:\Naziri-7-12-9.wiff
 Nominal mass (M_r): 66895; Calculated pI value: 5.19
 NCBI BLAST search of [gi|11066029](#) against nr
 Unformatted [sequence string](#) for pasting into other applications
 Taxonomy: [Clostridium difficile](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: 48%
 Matched peptides shown in **Bold Red**

1 MNISKQIVSL LTMTFLT VTL YGNTSNASTK DTLTGSGRWE TAIKISQAGW
 51 **TKSESAVLVN DNSIADALSA TPF**AKAKDAP ILLTQSNKLD SRTKAELKRL
 101 GVKNVYLIGG SIALSSEIEK **QLNAENINFE** RISGNSRYDT SLKLAEKLDLDR
 151 EKSISKIVVV NGEKGLADAV SVGAIAAQEN MPIILSDSEN GTEVADNFID
 201 SKDIAK**SYVI GGTYSISNSV** ERSLPNATRI AGSSRSETNA K**HIEEFYK**DT
 251 DIKNIYVTKD GTK**NKN**DLID SLAVGVLA AK NSSPIVL AGN KLDTTQKDV L
 301 NTKIIDK**VTQ IGGLGN**ENVV EDILDIQEET KYTVETIDEL NAAIKRADAN
 351 **DIKFKPEKE** KTINNSFSIE **TKKTV**FIELD GR YRQTITLD IPNGK**FNNYA**
 401 **EIEGGV**KLKN IKNESLVNKG SIQDLDIYDE NGCKIENESS GEIWFVTIVE
 451 EANDVYIVNS GDITK**ISNNS** SSTIIRNSGN IDTVTGK**KKEP AISGNKPKVN**
 501 DTEKETKAAR GLNPRVEACS VPK**KDYV**MIT IPNSPKDSRY KIYYR**VVYNK**
 551 **PYAMDVGDKI NIGEW**TVAPT DEEPFLEKAK NGCYVEAVEV NTSTKEVSRW
 601 GR TNATDDGF

Figure 3. 39. Result of Mascot search using data from mass spectrometry analysis of recombinant Cwp66
 The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 48% of the sequence of native Cwp66 could be recovered from the recombinant protein.

Match to: [gi|126700333](#) Score: 712
putative sortase B [Clostridium difficile 630]
 Found in search of F:\Naziri-7-12-9.wiff
 Nominal mass (M_r): 26762; Calculated pI value: 9.09
 NCBI BLAST search of [gi|126700333](#) against nr
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Clostridium difficile 630](#)
 Links to retrieve other entries containing this sequence from NCBI Entrez:
[gi|255307756](#) from [Clostridium difficile ATCC 43255](#)
[gi|115251770](#) from [Clostridium difficile 630](#)
 Fixed modifications: Carbamidomethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Sequence Coverage: 52%
 Matched peptides shown in **Bold Red**

1 MKKLYRIVIN IILVLVILYS GFNIYSKLT K YNHDTKISSE LQKKEYKKED
 51 LSKINSDFK**F WLSVENTNIN YPVVQSKDNS** YYLDKDFYKK DSISGTLFMD
 101 **YRNKSIDDKN IIIYGH**NMKN KTMFNLNK**F KDADFFK**KNN KIKITLNGKE
 151 FLYDVFSAYI VESDYDYLK**T NFN**NESDYQN YINDITSKSL YK**SPIKVNSN**
 201 **DKIVTLSTCT YEFDD**ARMVI HGRLI

Figure 3. 40. Result of Mascot search using data from mass spectrometry analysis of recombinant Sortase B.

The sequence in bold red text shows peptides identified in the analysis that match the sequence of the target protein from *C. difficile* 630. The data shows that 52% of the sequence of native Sortase b could be recovered from the recombinant protein.

3.3 Tomlinson phage display

Having successfully prepared recombinant target proteins in *E. coli*, the project moved to use these as targets for the isolation of recombinant scFv antibodies using phage display. The Tomlinson libraries I and J were used for phage display for several reasons. Firstly, they are highly diverse, estimated to contain around 10^8 unique specificities created through synthetic diversification of key residues in the CDRs of single human heavy and light chain scaffolds. Secondly, the libraries are well-suited to simple panning methods and the use of a trypsin-sensitive helper phage (KM13), combined with a protocol for trypsin elution of phage from the target, minimises the carry forward of virus with irrelevant properties. Finally, the antibodies that the libraries contain are scFvs, encoded on a phagemid vector. These can be easily and conveniently expressed and purified in high yield once target-specific scFvs have been isolated.

3.3.1 Helper phage preparation and titration

Phage KM13 provided with the Tomlinson library was used to prepare 100 fold serial dilutions. Each dilution was used to infect *E.coli* TG1 cells which were then plated in TYE H-top agar and incubated, allowing phage plaques to appear. To produce larger quantities of the helper phage, a small, clear plaque was picked and added into a fresh culture of *E.coli* TG1 that was growing in exponential phase. After incubation, the supernatant was recovered from the culture and was treated with PEG/NaCl to precipitate the KM13 phage. The pellet of viral particles was resuspended in PBS. This preparation was considered as a stock of helper phage and to determine titre, samples were titrated after trypsin treatment and without this treatment. In both, serial dilutions were prepared and infected to TG1 cells which were then plated in TYE H-top agar to allow plaque formation. Table 3.6 shows the difference observed in titration of trypsin-treated phage and untreated

phage. Firstly, the Table shows a high titre stock of helper phage had been prepared (about 4×10^{13} pfu/ml). Secondly, the data shows the sensitivity of KM13 to trypsin treatment as a difference in titre of about 5 logs was observed between treated and untreated samples. This KM13 helper phage was used as a stock in the later steps of phage display.

3.3.2 Preparation and titration phage from scFv libraries I and J:

The bacterial stocks of libraries I and J were grown in liquid medium and once exponential phase had been reached, samples were infected with KM13 helper phage. The remaining culture was used to prepare a secondary stock of both libraries. The infected cultures were grown on to allow the formation of phage displaying scFv on the surface. These were precipitated from culture supernatants using PEG/NaCl and stored in PBS ready for panning against target proteins. Each scFv phage library was titrated separately by serial dilution, infection of TG1 bacteria and plating on TYE-ampicillin plates. Table 3.7 shows the numbers of ampicillin resistant colonies recovered in this way from libraries I and J.

Table 3. 5 Titration of KM13 helper phage

NON-TRYPSIN TREATED PHAGE(PFU/ML)	TRYPSIN TREATED PHAGE(PFU/ML)	DIFFERENCE
3.69×10^{13}	1.15×10^8	2.54×10^5

Table 3. 6. Titration of scFv display I and J

LIBRARY I	LIBRARY J
8.1×10^{14}	4.8×10^{14}

3.3.3 Antibody selection from phage libraries I and J

Both libraries I and J contain more than 10^8 unique scFvs, and these resources were used to select recombinant antibodies with the capacity to bind the purified recombinant proteins described earlier and different forms of SlpA, the major surface layer protein of *C. difficile*. Selection of recombinant antibodies was carried out in two groups; in group one, libraries I and J were used separately for panning but in group two libraries I and J were mixed to reduce the complexities of the experiment. Group one screened for antibodies that would bind CspA, GroEL, FliD and the LMW form of SlpA as the targets. Group two comprised panning with the N-terminus of Cwp66, FliC, the C-terminus of Cwp66, Sortase B, Cwp66 and native SlpA extracted from *C. difficile*. To select the antibodies, immunotubes were coated with each target at 100µg/ml and panning was carried out as described below. In later rounds of selection, the concentration of each target was decreased to 75 and 50 µg/ml for the second and third rounds respectively.

Each round of selection comprised several steps: coating of the immunotube with the target; blocking unoccupied sites on the immunotube using skimmed milk in PBS; adding library phage and allowing target interaction; washing out unbound phage and elution of absorbed virus; amplification of the eluted phage for use in the next round of selection. In each round, TG1 cells were used to determine by titration the numbers of phage added to the immunotube (input) and the numbers of phage recovered at elution (output). Using data at input and output, the percentage recovery of phage in each round of selection was calculated by division of output by input titre. Tables 3.8 – 3.17 show the data for all targets over three rounds of selection.

Table 3. 7. Recovery of phage during the selection against CspA

LIBRARY	TOMLINSON I			TOMLINSON J		
	Input	Output	% Recovery	Input	Output	% Recovery
Round 1	8.1×10^{13}	1.6×10^5	1.9×10^{-7}	4.8×10^{13}	3.2×10^5	6.6×10^{-7}
Round 2	10^{13}	2.3×10^6	2.3×10^{-5}	3×10^{13}	1.3×10^6	4.3×10^{-6}
Round3	7.1×10^{13}	1.2×10^7	1.6×10^{-5}	3.8×10^{13}	3.2×10^8	8.4×10^{-4}

Table 3. 8. Recovery of phage during the selection against GroEL

LIBRARY	TOMLINSON I			TOMLINSON J		
	Input	Output	% Recovery	Input	Output	% Recovery
Round 1	8.1×10^{13}	1.6×10^5	1.9×10^{-7}	4.8×10^{13}	6.8×10^5	1.4×10^{-6}
Round 2	8.7×10^{13}	6.1×10^6	7×10^{-6}	5×10^{13}	7.2×10^6	1.4×10^{-5}
Round3	6.6×10^{13}	4.32×10^8	6.5×10^{-4}	8.3×10^{13}	4.2×10^8	5×10^{-4}

Table 3. 9. Recovery of phage during the selection against FliD

LIBRARY	TOMLINSON I			TOMLINSON J		
Selection	Input	Output	% Recovery	Input	Output	% Recovery
Round 1	8.1×10^{13}	8×10^4	9.8×10^{-8}	4.8×10^{13}	4×10^4	8.3×10^{-8}
Round 2	3×10^{13}	10^6	3.3×10^{-6}	2×10^{13}	1.3×10^6	6.5×10^{-6}
Round3	9.4×10^{13}	6.3×10^8	6.7×10^{-4}	8.8×10^{13}	1.06×10^9	1.2×10^{-3}

Table 3. 10. Recovery of phage during the selection against LMW

LIBRARY	TOMLINSON I			TOMLINSON J		
Selection	Input	Output	% Recovery	Input	Output	% Recovery
Round 1	8.1×10^{13}	1.2×10^4	1.4×10^{-8}	4.8×10^{13}	4×10^4	8.3×10^{-8}
Round 2	8.2×10^{13}	5.2×10^6	6.3×10^{-6}	5.4×10^{13}	7.2×10^6	1.6×10^{-5}
Round3	9×10^{13}	4.64×10^7	5.1×10^{-5}	9.7×10^{13}	8.3×10^8	8.5×10^{-4}

Table 3. 11. Recovery of phage during the selection against N-terminus of Cwp66

LIBRARY	TOMLINSON I+J		
Selection	Input	Output	% Recovery
Round 1	6.45×10^{13}	2.8×10^5	4.3×10^{-7}
Round 2	6×10^{13}	1.1×10^6	1.8×10^{-6}
Round3	1.3×10^{13}	6.6×10^7	1.1×10^{-4}

Table 3. 12. Recovery of phage during the selection against FliC

LIBRARY	TOMLINSON I+J		
Selection	Input	Output	% Recovery
Round 1	6.45×10^{13}	1.6×10^5	2.4×10^{-7}
Round 2	5×10^{13}	7.1×10^6	1.4×10^{-5}
Round3	8.1×10^{13}	1.06×10^7	1.3×10^{-5}

Table 3. 13. Recovery of phage during the selection against C-terminus of Cwp66

LIBRARY	TOMLINSON I+J		
Selection	Input	Output	% of Recovery
Round 1	6.45×10^{13}	2.4×10^5	3.7×10^{-7}
Round 2	9.7×10^{13}	2.1×10^6	2.1×10^{-6}
Round3	9.1×10^{13}	9.3×10^6	10^{-5}

Table 3. 14. Recovery of phage during the selection against Sortase B

LIBRARY	TOMLINSON I+J		
Selection	Input	Output	% Recovery
Round 1	6.45×10^{13}	1.4×10^6	2.1×10^{-6}
Round 2	10^{13}	4.4×10^4	4.4×10^{-7}
Round3	1.2×10^{13}	7.1×10^6	5.9×10^{-5}

Table 3. 15. Recovery of phage during the selection against Cwp66

LIBRARY	TOMLINSON I+J		
Selection	Input	Output	% Recovery
Round 1	6.45×10^{13}	8×10^5	1.2×10^{-6}
Round 2	3×10^{13}	7×10^6	2.3×10^{-5}
Round3	1.2×10^{13}	4.7×10^5	3.9×10^{-6}

Table 3. 16 Recovery of phage during the selection against SlpA

LIBRARY	TOMLINSON I+J		
Selection	Input	Output	% Recovery
Round 1	6.45×10^{13}	3.6×10^4	5.5×10^{-8}
Round 2	2×10^{13}	2.7×10^5	1.3×10^{-6}
Round3	1.1×10^{13}	1.6×10^5	1.4×10^{-6}

3.3.4 Antibody screening

3.3.4.1 Monoclonal phage ELISA

Whilst the recoveries of virus from panning suggested enrichment of target-specific scFvs, this could only be confirmed by immunoassay. To do this, individual colonies from the 3rd round of each selection were picked for analysis. In the group one screens – those in which libraries I and J were used separately for selection against CspA, GroEL, FliD and LMW of SlpA – 48 colonies were picked from each library in the final selection step. For group two screens in which the two libraries were combined for selection, a total of 48 colonies were picked. Group two screens panned for scFvs against the N-terminus of Cwp66, FliC, the C-terminus of Cwp66, Sortase B, Cwp66 and native SlpA.

Individual colonies were picked at random from the third round of selection against each target and grown overnight in 96 well plates. Cultures were re-inoculated to medium, grown into exponential phase and infected with helper phage. Having thereby generated monoclonal phage particles displaying scFvs, they were then transferred to ELISA plates pre-coated with the relevant target protein. The binding of phage to the targets was detected with anti-M13-HRP reagents. Figures 3.41 – 3.47 illustrate the ability of individual phage clones to recognise and bind to the selecting target.

In Figures 3.41 to 3.44, phage ELISA revealed that 50-100% of all clones were reactive against the intended target although there was some variation in the signal strength as measured by absorbance at 450 nm. Figure 3.41 shows that more than 60% of all selected colonies from both libraries panned against CspA were positive in phage ELISA. The strongest reactions against CspA came from clones from library I as measured by absorbance at 450 nm. Higher reaction frequency can be seen in Figure 3.42 in which more than 90% the clones picked from selection against GroEL were positive in phage ELISA. In this experiment, the

strongest reaction came from clones from library J. An interesting contrast is evident in Figure 3.43 in which higher frequency of reaction with FliD arose from selection using library J than library I. The strongest reactions in this ELISA also came from clones from library J. The same pattern emerged from experiments with LMW (Figure 3.44), with 90% of clones picked for analysis showing strong positive reaction for the target, particularly from library J, although strong reactions can be also seen for clones from library I.

Figures 3.45 to 3.47 show the properties of phage recovered from selections with a mixture of libraries I and J. Figures 3.45 and 3.47 show contrasting results from screens using isolated regions of Cwp66 (N and C terminal regions, Figure 3.45) and the intact Cwp66 protein (Figure 3.47, left panel). Nearly all clones picked from screens against N and C-terminal part of Cwp66 are positive with a high absorbances in phage ELISA (Figure 3.45). In contrast, only 30% of clones isolated against Cwp66 were positive in phage ELISA, just half of them with high signal strength (Figure 3.47, left panel). Figure 3.46 shows identification of target-specific clones in phage ELISA against FliC (left panel) and Sortase B (right panel). For all targets, phage ELISA identified ample numbers of positive clones for the next steps of the project, however to maximise the chances of success, all clones identified as showing target-specificity in phage ELISA were taken forward.

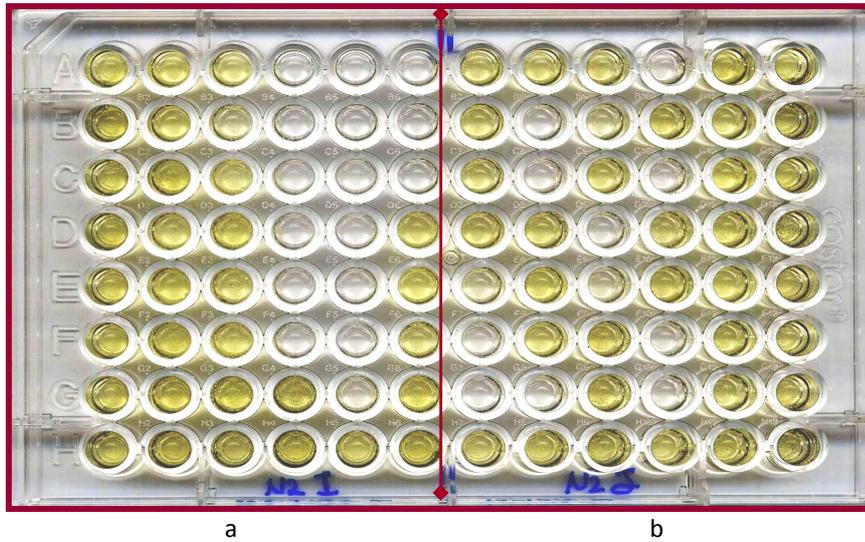


Figure 3. 41. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant CspA
Binding of individual phage particles from the third round of selection was detected using anti-M13-HRP and TMB as the enzyme substrate

a) 48 colonies picked randomly from 3rd round of selection using of library I

b) 48 colonies picked randomly from 3rd round of selection using of library J

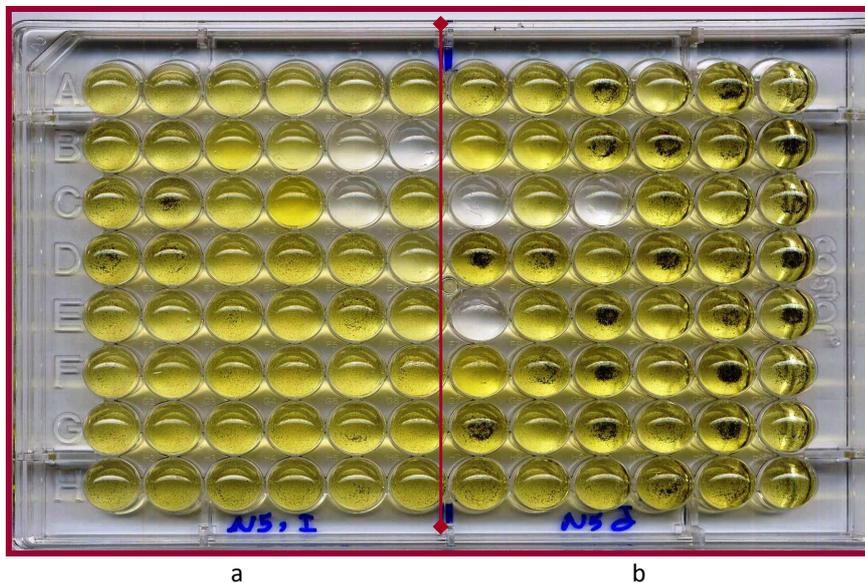


Figure 3. 42. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant GroEL

Binding of individual phage particles from the third round of selection was detected using anti-M13-HRP and TMB as the enzyme substrate

a) 48 colonies picked randomly from 3rd round of selection using of library I

b) 48 colonies picked randomly from 3rd round of selection using of library J

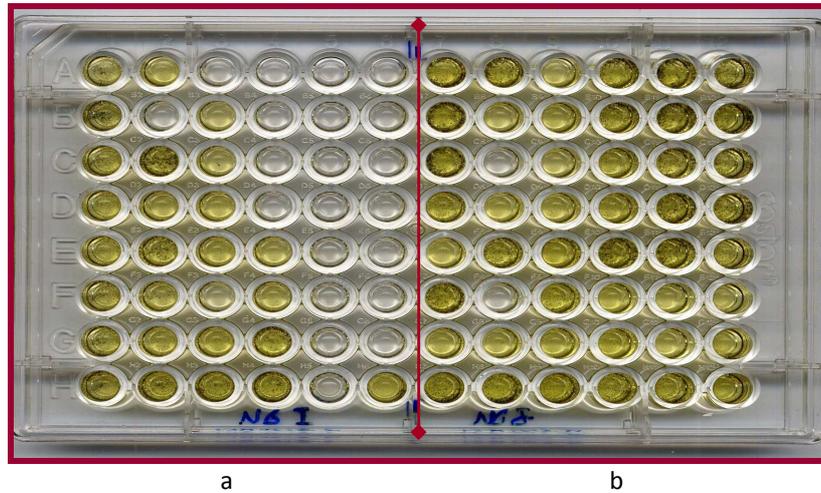


Figure 3. 43. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant FliD
Binding of individual phage particles from the third round of selection was detected using anti-M13-HRP and TMB as the enzyme substrate

- a) 48 colonies picked randomly from 3rd round of selection using of library I
b) 48 colonies picked randomly from 3rd round of selection using of library J

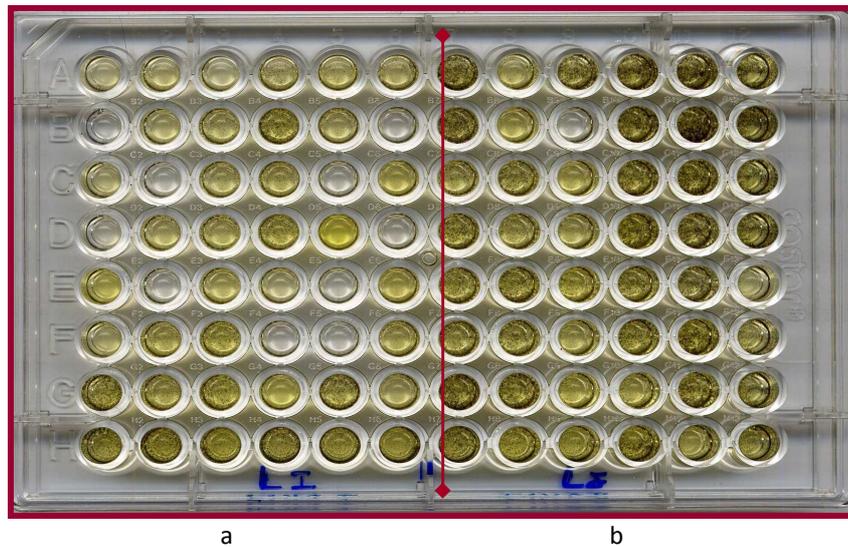


Figure 3. 44. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant LMW

Individual phage particles from third round of selection from Libraries I and J in a ELISA test using antiM13-HRP and TMB as the enzyme substrate

- a) 48 colonies picked randomly from 3rd round of selection of library I
b) 48 colonies picked randomly from 3rd round of selection of library J

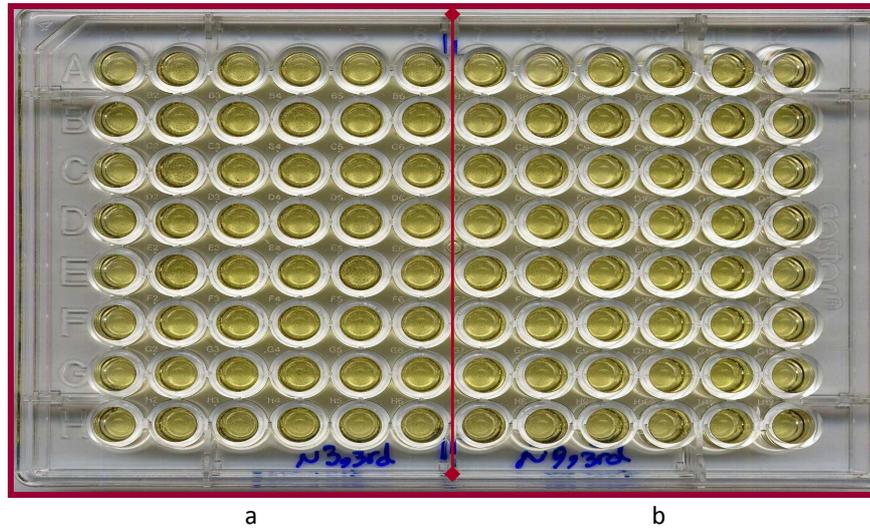


Figure 3. 45. Monoclonal phage ELISA of phage from Tomlinson libraries I+J against recombinant N-terminus (left) and C-terminus (right) of Cwp66

Binding of individual phage particles from the third round of selection was detected using anti-M13-HRP and TMB as the enzyme substrate

- a) 48 colonies picked randomly from 3rd round of selection using of libraries I+J
- b) 48 colonies picked randomly from 3rd round of selection using of libraries I+J

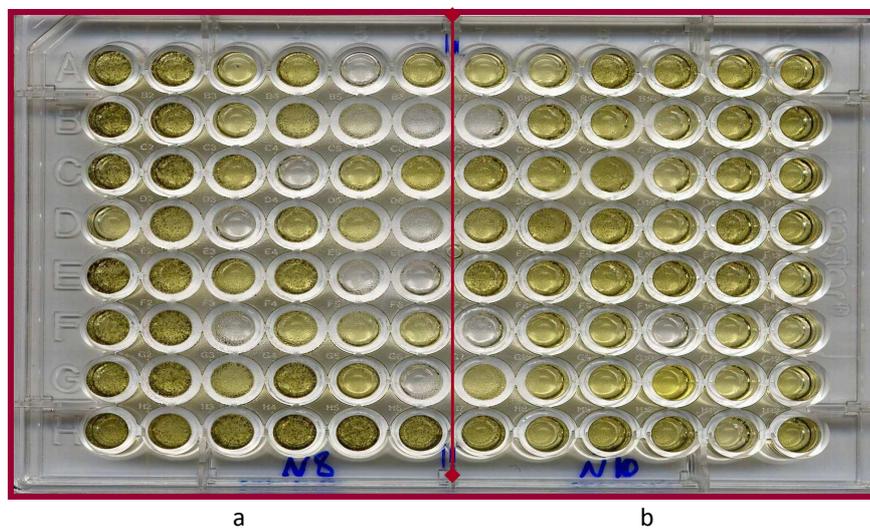


Figure 3. 46. Monoclonal phage ELISA of phage from Tomlinson libraries I+ J against recombinant FliC (left) and Sortase B (right)

Binding of individual phage particles from the third round of selection was detected using anti-M13-HRP and TMB as the enzyme substrate

- a) 48 colonies picked randomly from 3rd round of selection using of library I+J
- b) 48 colonies picked randomly from 3rd round of selection using of library I+J

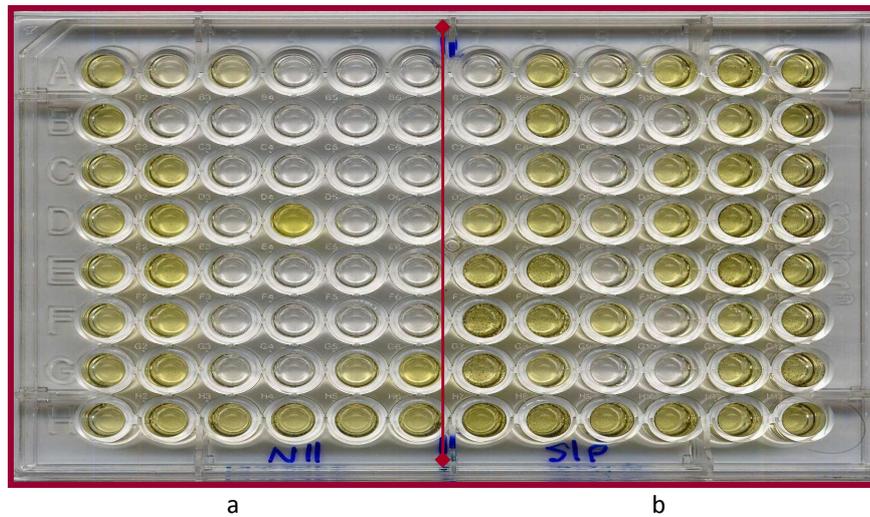


Figure 3. 47. Monoclonal phage ELISA of phage from Tomlinson libraries I and J against recombinant Cwp66 (left) and extracted SlpA (right)

Binding of individual phage particles from the third round of selection was detected using anti-M13-HRP and TMB as the enzyme substrate

a) 48 colonies picked randomly from 3rd round of selection using of library I+J

b) 48 colonies picked randomly from 3rd round of selection using of library I+J

3.3.4.2 Screening of monoclonal soluble recombinant antibodies

In phage display using the Tomlinson libraries and related systems, phage are produced in *E. coli* TG1, a strain that suppresses the TAG stop codon at the junction between the recombinant scFv antibody and gIII, the gene for a minor phage coat protein pIII. This fuses the scFv to the coat protein so that the scFv is displayed at the surface of the virus. TAG stop codons present in the scFv reading frame will be similarly suppressed and hence the next phase of analysis was to express soluble scFv (*ie* unfused to pIII, achieved by expression in a non-suppressor strain of *E. coli*) to see if target reactivity was retained.

Also, it is likely that binding of phage to the target is a multivalent interaction. Five copies of pIII are present at the phage surface and whilst it is unlikely that all 5 will be scFv-pIII fusions, scFvs may be isolated that possess very low affinity for the target and only bind during panning and phage ELISA as a result of multivalent scFv display. To exclude scFvs of this sort, the antibodies were expressed as soluble scFvs in *E. coli* HB2151 (a non-suppressor strain) thereby restricting them to monovalent interaction with target in ELISA. The scFvs can be detected in ELISA because they retain the c-myc detection tag (Figure 3.65) and can be purified on a large scale because of the histidine repeats at the C-terminus of the scFv sequence.

Soluble scFv proteins were prepared from all clones identified as positive in monoclonal phage ELISA. Initially, phage were prepared from each clone by growing the bacteria into exponential phase, infecting with helper phage and growing on overnight. These phage antibodies from the supernatant of overnight cultures were used to infect *E. coli* HB2151 cells; successfully infected HB2151 cells were selected by plating on medium containing ampicillin. New liquid cultures were then established and expression of soluble scFv was induced by addition of

IPTG. Culture supernatants were first tested for the presence of soluble scFv by dot blotting. The ability of scFv to bind to the target proteins was then assessed by ELISA. To speed throughput and ease record keeping, these experiments were carried out in 96 well plates.

3.3.4.2.1 Dot blot analysis for soluble recombinant antibodies

Depending on the number of clones to be screened, a grid was drawn out on nitrocellulose membrane and samples of each culture supernatant were added. As soluble scFvs retain the c-myc tag towards the C-terminus of the scFv protein, the presence of recombinant antibody could be detected with anti c-myc and anti-rabbit-HRP reagents. Data is shown in Figures 3.48 to 3.54. Although the Figures clearly show that a positive response was obtained for many, if not all the supernatants that were tested, the intensity of signals are different which might reflect differing scFv yields from the clones under test or variable volumes of culture supernatant. In all figures, a wide variation in signal intensity is evident, with variable frequencies of positive reaction.

For instance in Figure 3.48 many positive signals can be seen but two highly intense reactions are evident (A4 and G12), one clone from each of the Tomlinson libraries. The signals for soluble scFvs against GroEL (Figure 3.49) are generally stronger and there is greater representation of high signals from libraries I and J (Figure 3.49). Dot blot analysis of soluble scFvs against FliD in Figure 3.50 reveals a marked difference between the two libraries with many more positive signals from library I than library J, a pattern that is also evident in Figure 3.51 (soluble antibodies against LMW).

Results from Figure 3.52 to 3.54 are from screens in which the two libraries were mixed at the start of the panning process. Strong positive reactions can be

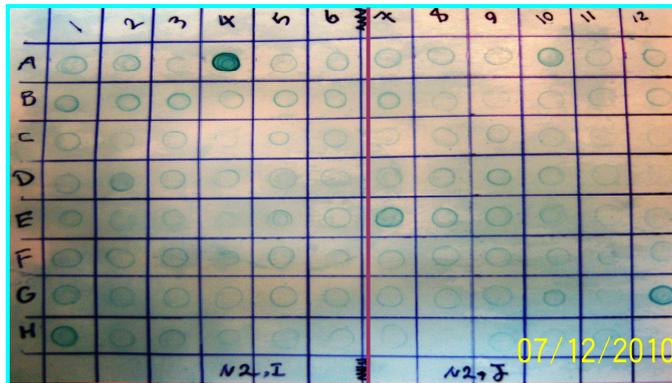
seen in all cases though in one (Figure 3.53, right panel, anti-Sortase B antibodies) only a single strong signal appeared.

3.3.4.2.2 Detection of soluble antibodies by ELISA

The foregoing experiments tested only for the production of scFv proteins. To test for target specificity, the supernatants of induced HB2151 cultures were used in immunoassays against the recombinant target proteins. Binding of scFv to target-coated ELISA plates was detected with anti-c-myc and anti rabbit-HRP reagents and signals were quantified by absorbance at 450 nm. Results are shown in Figures 3.55 to 3.62. Results from Figure 3.55 show most of the antibodies from library J against recombinant CspA (right part of the plate, b) have a strong binding reaction when compared to clones from library I (left part of the plate, a).

The reaction of selected monoclonal antibodies against GroEL in Figure 3.56 revealed far fewer positive clones but the best were derived from library J (right part of the plate, b). Figure 3.57 shows results with FliC (left panel, a) and SortaseB (right panel, b). While more than 90% of antibodies against Sortase B could react with their target, this decreased to around 40% for the binders selected against FliC. Monoclonal soluble scFv against LMW were selected from both libraries as can be seen in Figure 3.58. The data confirmed that strong binders were isolated from both libraries, but the frequency was greater from library J (right panel, b). In Figure 3.59, the majority of scFvs against the N-terminus of Cwp66 were able to bind to the target, a finding that was paralleled for the scFvs against C-terminus of Cwp66 (Figure 3.61, right panel). In contrast, only modest numbers of monoclonal antibodies were reactive with Cwp66 (around 25% of the 48 clones picked for analysis; Figure 3.61, left panel). Almost all clones selected from library I against FliD were reactive in soluble scFv ELISA (Figure 3.60, left panel) whereas about 50%

of clones from library J were positive (Figure 3.60, right panel). Figure 3.62 shows that the majority of the soluble scFvs against SlpA were reactive with the target.



a

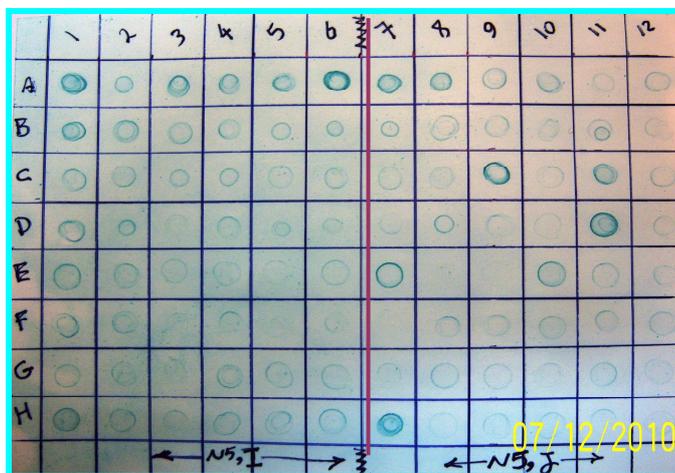
b

Figure 3. 48. Dot blot analysis of expression of soluble scFv against recombinant CspA

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection using library I

(b) Supernatants derived from the 3rd round of selection using library J



a

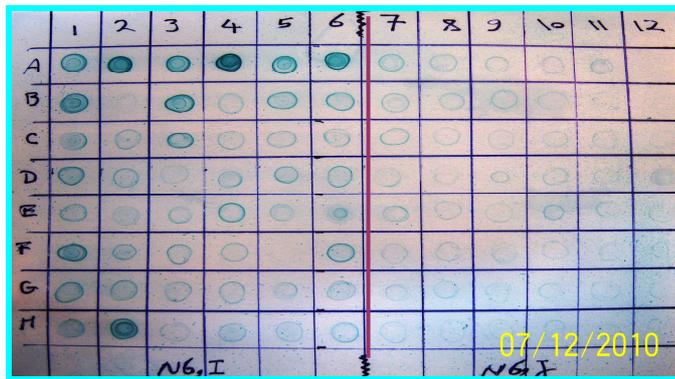
b

Figure 3. 49. Dot blot analysis of expression of soluble scFv against recombinant GroEL

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection using library I

(b) Supernatants derived from the 3rd round of selection using library J



a

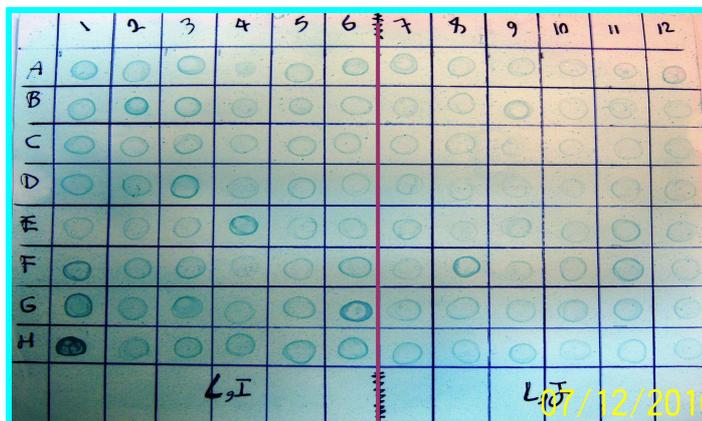
b

Figure 3. 50. Dot blot analysis of expression of soluble scFv against recombinant FliC

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection using library I

(b) Supernatants derived from the 3rd round of selection using library J



a

b

Figure 3. 51. Dot blot analysis of expression of soluble scFv against recombinant LMW

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection using library I

(b) Supernatants derived from the 3rd round of selection using library J

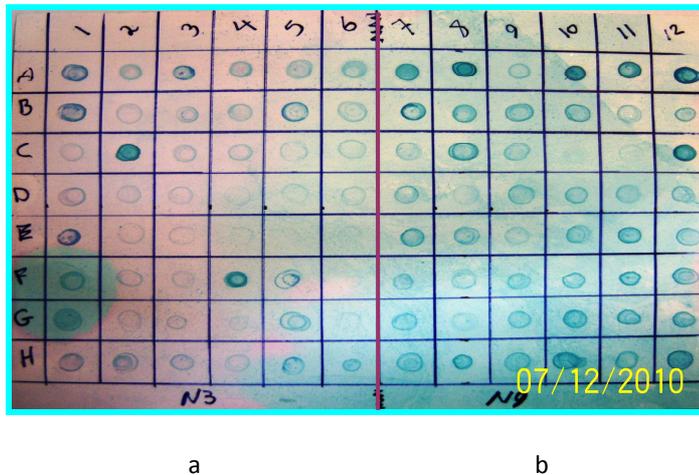


Figure 3. 52. Dot blot analysis of expression of soluble scFv against recombinant N-terminus and C-terminus of Cwp66

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection against the N-terminus of Cwp 66 using library I+J

(b) Supernatants derived from the 3rd round of selection against the C-terminus of Cwp 66 using library I+J

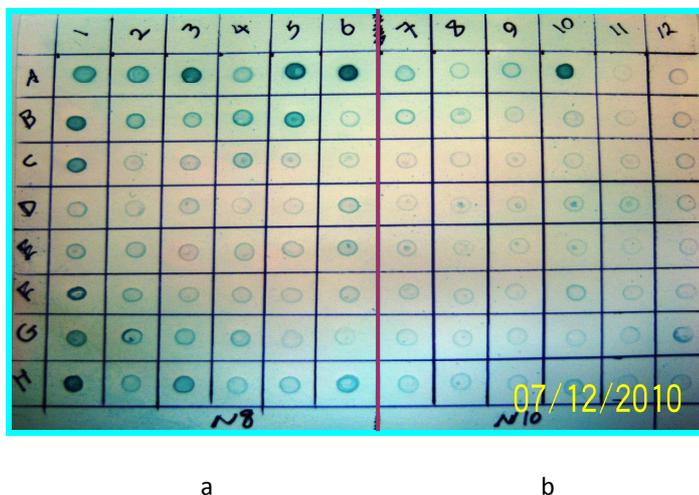
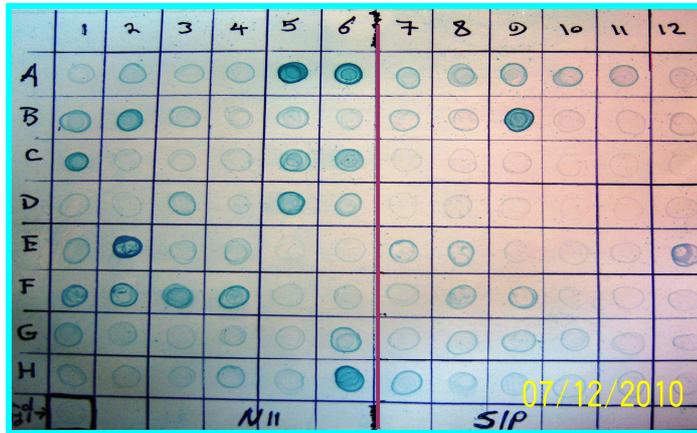


Figure 3. 53. Dot blot analysis of expression of soluble scFv against recombinant FliC and Sortase B

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection against FliC using library I+J

(b) Supernatants derived from the 3rd round of selection against Sortase B using library I+J



a

b

Figure 3. 54. Dot blot analysis of expression of soluble scFv against recombinant Cwp66 and extracted SlpA

Supernatants from induced cultures of *E. coli* HB2151 were blotted on Hybond membrane and scFv detected with reagents to the c-myc tag.

(a) Supernatants derived from the 3rd round of selection against Cwp66 using library I+J

(b) Supernatants derived from the 3rd round of selection against SlpA using library I+J

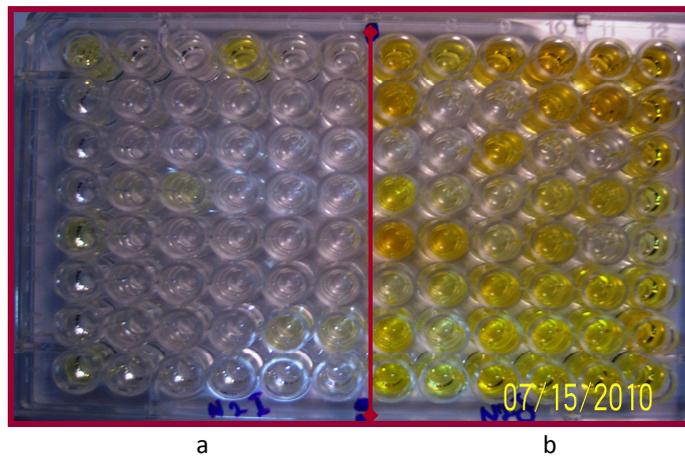


Figure 3. 55. Monoclonal soluble scFv ELISA against recombinant CspA

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant CspA. After adding scFvs to the coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

(a) scFvs from 3rd round of selection of library I against recombinant CspA

(b) scFvs from 3rd round of selection of library J against recombinant CspA

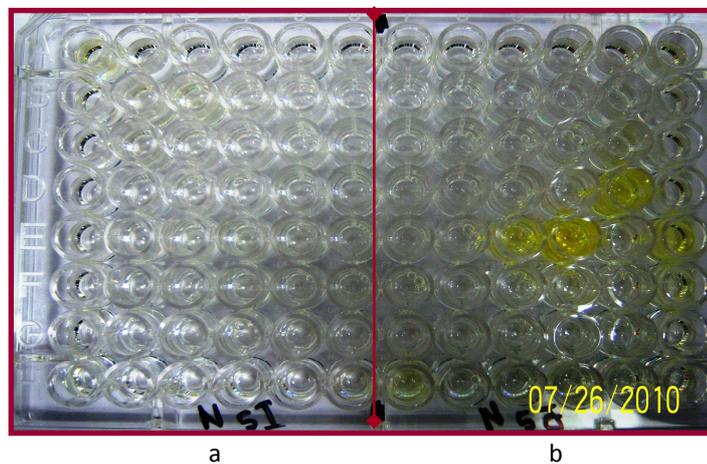


Figure 3. 56 Monoclonal soluble scFv ELISA against recombinant GroEL

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant GroEL. After adding scFvs to the coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

(a) scFvs from 3rd round of selection of library I against recombinant GroEL

(b) scFvs from 3rd round of selection of library J against recombinant GroEL

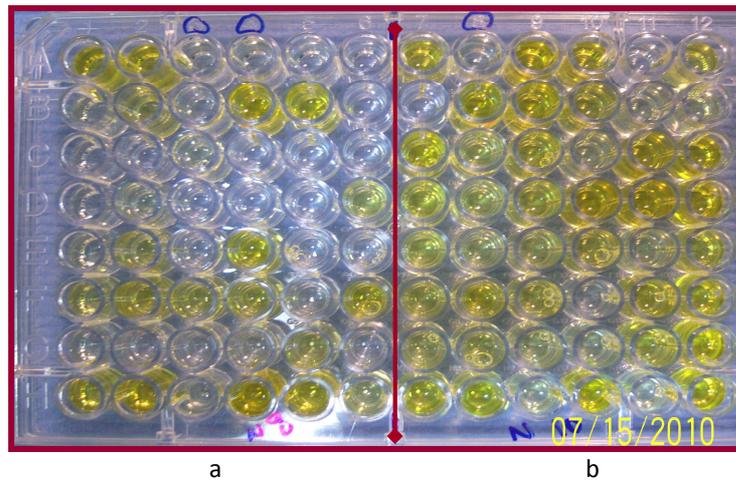


Figure 3. 57. Monoclonal soluble scFv ELISA against recombinant FliC and Sortase B

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant FliC and Sortase B. After adding scFvs to the coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

(a) scFvs from 3rd round of selection of libraries I+J against recombinant FliC

(b) scFvs from 3rd round of selection of libraries I+J against recombinant Sortase B

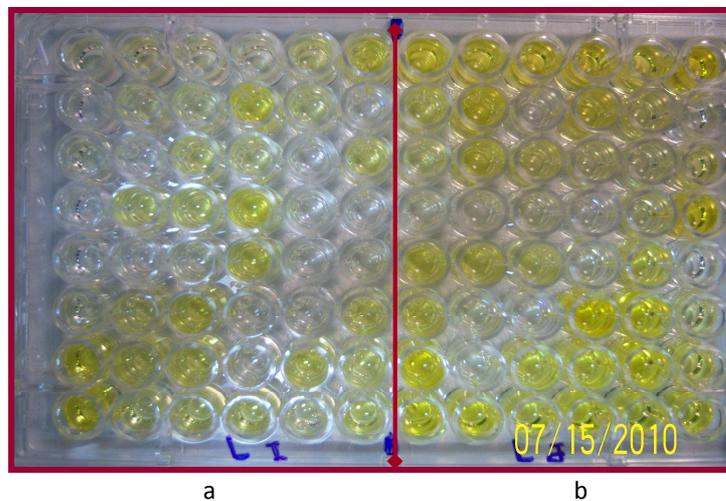


Figure 3. 58. Monoclonal soluble scFv ELISA against recombinant LMW

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant LMW. After adding scFvs to coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

(a) scFvs from 3rd round of selection of library I against recombinant LMW

(b) scFvs from 3rd round of selection of library J against recombinant LMW

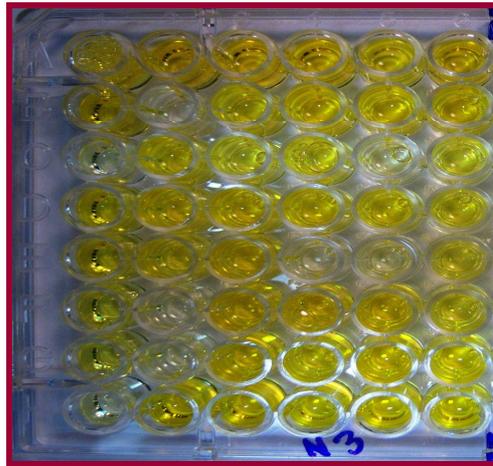


Figure 3. 59. Monoclonal soluble scFv ELISA against recombinant N-terminus Cwp66

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant N-terminus Cwp66. After adding scFvs to coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

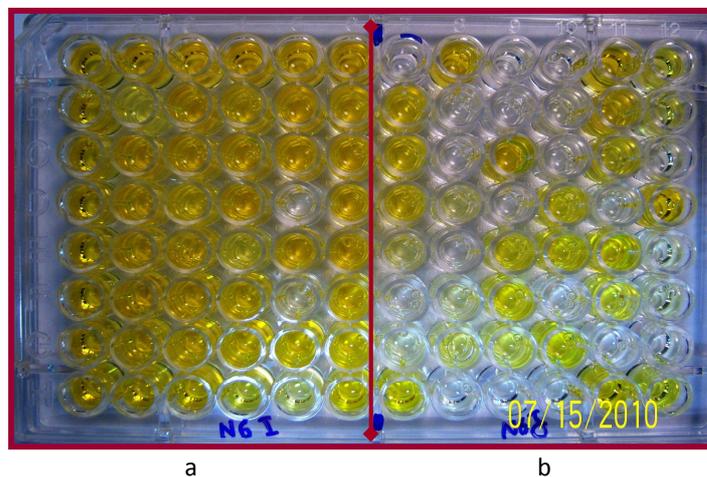


Figure 3. 60. Monoclonal soluble scFv ELISA against recombinant FlID

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant FlID. After adding scFvs to coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

(a) scFvs from 3rd round of selection of library I against recombinant FlID

(b) scFvs from 3rd round of selection of library J against recombinant FlID

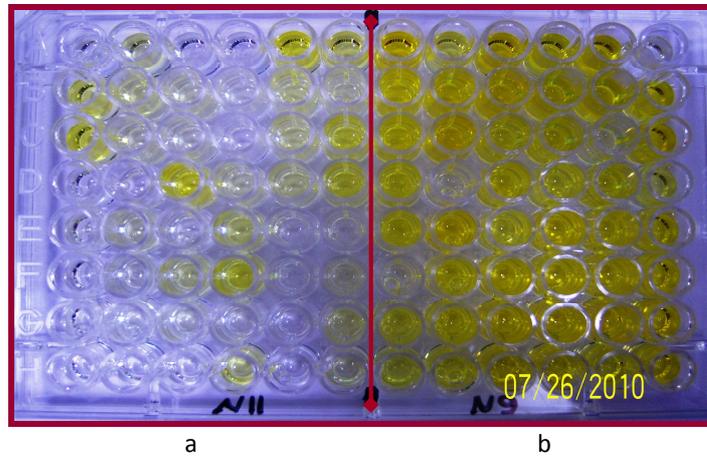


Figure 3. 61. Monoclonal soluble scFv ELISA against recombinant Cwp66 and the C-terminal region of the protein

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified recombinant C-terminus Cwp66 and Cwp66. After adding scFvs to coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

(a) scFvs from 3rd round of selection of libraries I+J against recombinant Cwp66

(b) scFvs from 3rd round of selection of library J against recombinant C-terminus Cwp66



Figure 3. 62. Monoclonal soluble scFv ELISA against extracted SlpA

Recombinant soluble antibodies from 3rd round of selection were detected in ELISA against purified extracted SlpA. After adding scFvs to coated, blocked plate, antibody binding was detected with anti-c-myc and anti-rabbit-HRP reagents and substrate.

3.3.5 Characterisation of selected clones

3.3.5.1 Sequence analysis of selected scFvs

In total, 672 clones were picked against recombinant target proteins and tested by phage ELISA, monoclonal phage ELISA and other immunoassays. Based on the monoclonal screening described earlier and Western blot analysis (subsequent sections), 32 scFv clones were chosen which possessed high binding ability to their targets. To reveal their amino acid sequences – in particular the sequences of the CDRs – plasmid DNA was prepared from all 32 HB2151 clones for sequencing with primers LMB3 and gIII (Table 3.17). Sequencing with these primers determined the full sequence of the scFv reading frame including VH, linker and VL. All sequences were checked for the restriction sites present in the pIT2 vector (Figure 3.65) and aligned to VBASE sequences using DNAPLOT database (<http://www.vbase2.org/vbase2.php>) to predict their amino acid sequences. Predicted amino acid sequences were aligned with amino acid sequences of the single framework sequences of the Tomlinson libraries (VH: V3-23/DP-47 and VL: O12/DPK9) to reveal the CDRs .

The Tomlinson libraries are based on single human VH and VL frameworks with invariable CDR1 and diversity at defined points in CDR2 and CDR3. Specifically, the scFv heavy chain domain is based on the V3-23/DP-47 sequence and diversified in CDR2 (residues H50, H52, H52a, H53, H55, H56, and H58) and CDR3 (residues H95, H96, H97 and H98). In the light chain, the O12/O2/DPK9 sequence is diversified in CDR2 (residues L50 and L53) and CDR3 (L91 L92, L93, L94, L96).

Table 3.18 shows the basic amino acid sequences of heavy and light chains of each scFv, connected into a scFv by a linker sequence. The diversified amino acid sequences of the sequenced clones for both heavy and light chains were extracted (Tables 3.19 [heavy chains] and 3.20 [light chains]).

In limited instances, clones were identical in sequence but the more common finding was that scFvs against particular target proteins shared general properties in CDRs 2 and 3. For example, the three antibodies against CspA were all unique in sequence through the VH and VL domains but the heavy chain CDR2 sequences all possess two or more hydroxylated residues (N2A9, one threonine, one serine; N2E7, five threonines; N2G12, one threonine, two serines), and in CDR3, 2 of the 4 diversified residues were lysine or arginine. Taking scFvs against the N-terminus of Cwp66, clones N3A1 and N3H2 were closely related in sequence whereas the third sequenced clone (N3F4) was quite different. N3A1 and N3H2 shared an xGYSSAx sequence in CDR2 in the heavy chain, VH CDR3 sequences were identical and the light chain sequences were identical throughout.

N11 clones were directed against the complete Cwp66 protein. Sequences were identical in VH at the diversified amino acids except just one (H50 of N11F4). In the sequence of VH for the N11 clones there are hydroxylated amino acids at one side of CDR2 (one threonine and one serine) and hydrophobic amino acid (isoleucine) at the middle and end of this region. The nature of CDR3 residues of VH of these clones are mostly hydrophobic. In VL sequences of the N11 clones, two are identical (C1 and F4) with a mix of hydrophobic and hydroxylated amino acids at CDR3. Clone N11A5 is notably for the presence of tryptophan in the CDR3 of the VL domain; no other VL sequence possessed this feature and there was only one example amongst the VH sequences (N5E10; a scFv against GroEL). scFvs against the N-terminus of Cwp66 were similar but non-identical, a feature that also applied to anti-FliD scFvs (N8 clones).

scFvs against SlpA were isolated by selection specifically against the LMW component and by panning against cell surface extracts. The six clones against LMW could be subdivided: LA6 and LD4 were identical and LG1 and LH1 formed a second identical pair; LF10 and LG7 were unique. Although there were similarities amongst the anti-SlpA scFvs, the sequences were all unique.

The scFvs against FliC (N6 designation) are dissimilar in the CDRs of both the heavy and light chains, but an identical motif NYT (asparagine, tyrosine, threonine) can be seen in CDR2 of the VH for all antibodies. Just one clone against GroEL was analysed (N5E10). This generally carried charged residues at diversified points in CDR3 of VH and VL and uniformly hydrophobic residues at diversified points in CDR2 of VL. CDRs of the three anti-Sortase B scFvs (N10 clones) were very similar or identical.

3.3.5.2 Restriction analysis of selected clones

The Tomlinson libraries are based upon the pIT2 vector (Figure 3.65) in which a scFv is separated from gIII of the M13 phage by a histidine purification tag, the c-myc epitope and an amber stop codon. Four restriction sites (*NcoI*, *XhoI*, *Sall* and *NotI*) mark the boundaries of the VH and VL coding sequences. Digestion at these sites can be used to analyse whether individual clones carry a full length scFv or a truncated sequence. Restriction analysis was applied to verify whether the selected scFvs were indeed full length in sequence.

Fourteen selected clones in *E. coli* HB2151 were cultured for plasmid preparation. Purified plasmids were digested using three combinations of restriction enzymes. To check for the presence of a full length scFv, plasmids were digested with *NcoI* and *NotI*; Figure 3.66 illustrates the presence of these sites at the termini of the scFv insert. In the second and third digests, plasmids were treated with combinations of restriction enzymes to assess for the presence of VH and VL components. *NcoI* and *XhoI* were used to check for VH and digestion with *Sall* and *NotI* verified the presence of VL. Figure 3.67 shows that digestion of

plasmids with *NcoI* and *NotI* excised a DNA fragment of the same size from all clones. The size of the fragment was consistent with the presence of a full-length scFv insert in all clones. This was substantiated by analysis for the presence of VH and VL components (Figure 3.68 upper and lower sections respectively).

Table 3. 17. Forward (LMB3) and reverse (gIII) primers used in sequencing of final selected clones.

PRIMERS	SEQUENCE FROM 5' TO 3'
LMB3	CAGGAAACAGCTATGAC
gIII	CCCTCATAGTTAGCGTAACG
Lin Seq new	CGACCCGCCACCGCCCGCTG
DPK9 FR1 seq	CATCTGTAGGAGACAGAGTC
pHEN seq	CTATGCGGCCCATCA
PeIB	ATGAAATACCTATTGCCTACGGCAGC

Table 3. 18. Basic sequence of heavy and light chains of scFv which were diversified in the creation of the Tomlinson libraries.

	AMINO ACID SEQUENCE
Heavy chain V3-23/DP-47	EVQLLESGGGLVQPGGSLRLSCAAS GETESSY AMSWVRQAPGK GLEWWSAISGSGGSTYYADSVKGRFTISRDN SKNTLYLQMNSLR AEDTAVYYCAK
Light chain O12/DPK9	DIQMTQSPSSLSASVGDRTITCRAS QSISY LNWYQQKPGKAP KLLIYAASSLQSGVPSRFSGSGSGTDFLTISLQPEDFATYYCQ QSYSTPP
Linker	GGGGSGGGSGGGG

The highlighted sequences show CDR1 of both chains which are constant in all antibodies. The third row is the Gly-Ser linker sequence that connects VH and VL domains.

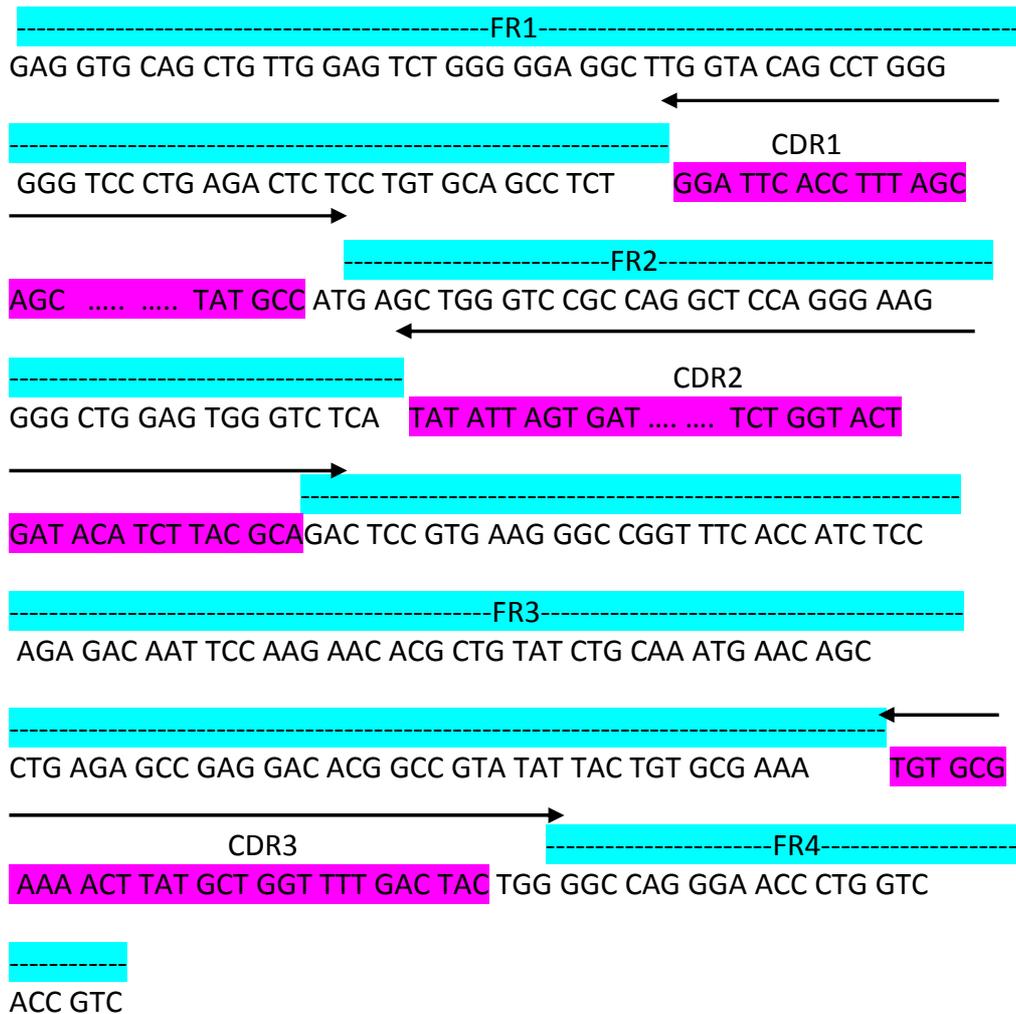


Figure 3. 63. Complete heavy chain nucleic acid sequence with framework regions (blue) and CDRs (red) highlighted

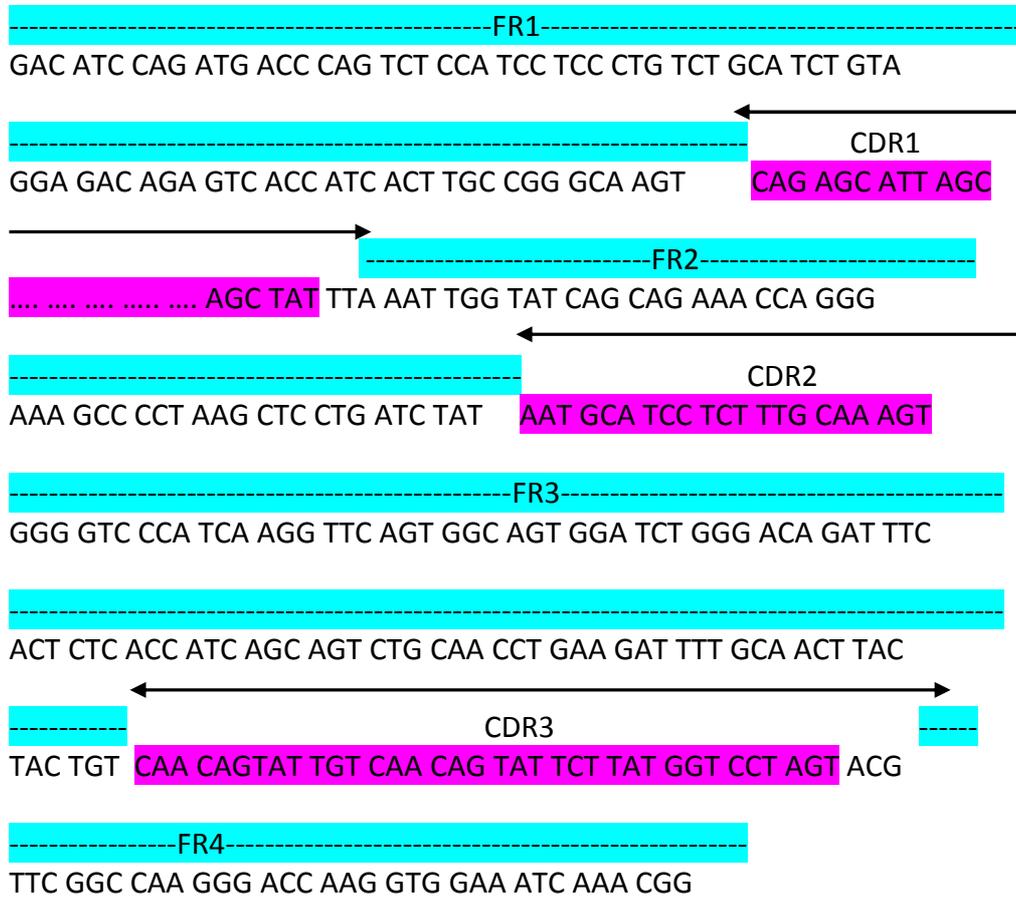


Figure 3. 64. Light chain nucleotide sequence with framework regions (blue) and CDRs (red) highlighted

Table 3. 19. Diversified amino acids in CDR2 and CDR3 in heavy chains of sequenced clones. Target specificities: N2, CspA; N3, N-terminus of Cwp66; N5, GroEL; N6, FliC; N8, FliD; N9, C-terminus of Cwp66; N10, SortaseB; N11, Cwp66, L, LMW; slp, SlpA

	CDR2						CDR3				
CLONES	H50	H52	H52A	H53	H55	H56	H58	H95	H96	H97	H98
N2A9	T	E	K	Q	E	S	D	R	K	P	P
N2E7	T	T	T	P	A	T	T	T	R	R	M
N2G12	S	S	R	T	K	G	F	L	S	K	R
N3A1	A	G	Y	S	S	A	R	N	A	Y	T
N3F4	T	S	R	L	T	N	L	N	G	A	L
N3H2	S	G	Y	S	S	A	I	N	A	Y	T
-----	----	---	---	---	---	---	---	---	---	---	---
N9A11	G	N	Y	D	Y	T	S	Y	S	T	S
N9B9	G	T	A	N	Y	T	S	T	G	S	N
N9C9	R	I	P	N	Y	T	E	S	G	H	T
-----	----	---	---	---	---	---	---	---	---	---	---
N11A5	T	S	R	L	T	N	L	N	G	A	L
N11C1	T	S	R	L	T	N	L	N	G	A	L
N11F4	S	S	R	L	T	N	L	N	G	A	L
-----	----	---	---	---	---	---	---	---	---	---	---
N5E10	H	T	K	E	T	G	W	R	R	H	H
-----	----	---	---	---	---	---	---	---	---	---	---
N6A4	G	N	S	N	Y	T	S	Y	G	S	N
N6B5	G	T	A	N	Y	T	S	S	S	T	S
N6C9	R	I	P	N	T	T	E	S	G	H	T
-----	----	---	---	---	---	---	---	---	---	---	---
N8B4	S	T	S	T	D	Y	G	T	S	S	N
N8F6	S	A	Y	G	D	Y	S	S	A	Y	N
-----	----	---	---	---	---	---	---	---	---	---	---
N10B8	T	T	S	L	D	F	S	P	A	A	T
N10D10	S	T	S	L	D	F	S	P	A	A	T
N10D11	S	T	S	L	D	F	Q	P	S	A	T

Table 3.19 continued

CLONES	CDR2						CDR3				
	H50	H52	H52A	H53	H55	H56	H58	H95	H96	H97	H98
LA6	S	S	G	T	Y	S	A	G	D	S	F
LD4	S	S	G	T	Y	S	A	G	D	S	F
LF10	S	Y	A	?	T	R	K	H	P	L	I
LG1	T	G	T	Y	S	S	A	N	A	A	A
LG7	G	S	T	H	S	R	Q	N	G	T	L
LH1	T	G	T	Y	S	S	A	N	A	A	A
-----	----	---	---	---	---	---	---	---	---	---	---
slpA9	G	S	T	H	S	R	Q	N	G	T	L
slpA10	T	S	T	A	K	G	H	N	Y	P	A
slpE10	L	N	A	A	T	R	L	S	M	R	A

Table 3. 20. Diversified amino acids in CDR2 and CDR3 in light chains of sequenced clones. Target specificities: N2, CspA; N3, N-terminus of Cwp66; N5, GroEL; N6, FliC; N8, FliD; N9, C-terminus of Cwp66; N10, SortaseB; N11, Cwp66, L, LMW; slp, SlpA

CLONES	CDR2			CDR3			
	L50	L53	L91	L92	L93	L94	L96
N2A9	R	H	R	M	R	A	L
N2E7	H	R	G	T	T	L	M
N2G12	A	S	Q	H	R	A	H
-----	---	---	---	---	---	---	---
N3A1	A	Y	S	D	N	S	A
N3F4	A	L	S	S	V	S	P
N3H2	A	Y	S	D	N	S	A
-----	---	---	---	---	---	---	---
N9A11	H	G	T	A	D	S	S
N9B9	N	S	D	G	N	S	D
N9C9	N	I	N	S	D	A	F
-----	---	---	---	---	---	---	---
N11A5	G	S	A	W	A	L	N
N11C1	A	L	S	S	V	S	P
N11F4	A	L	S	S	V	S	P
-----	---	---	---	---	---	---	---
N5E10	G	Y	F	S	R	K	S
-----	---	---	---	---	---	---	---
N6A4	D	N	S	N	S	Y	S
N6B5	N	Y	G	N	N	N	S
N6C9	Q	N	P	V	S	A	P
-----	---	---	---	---	---	---	---
N8B4	G	S	D	S	N	T	A
N8F6	S	A	A	S	N	T	G
-----	---	---	---	---	---	---	---
N10B8	A	S	T	T	S	T	D
N10D10	A	S	N	L	S	T	D
N10D11	A	S	N	T	S	T	D

Table 3.20 continued

CLONES	CDR2			CDR3			
	L50	L53	L91	L92	L93	L94	L96
LA6	S	T	A	S	S	S	T
LD4	S	T	A	S	S	S	T
LF10	M	M	K	K	A	T	A
LG1	A	S	Y	A	Y	Y	T
LG7	Q	E	A	Q	S	N	Q
LH1	A	S	Y	A	Y	Y	T
-----	---	---	---	---	---	---	---
slpA9	Q	E	A	Q	S	N	Q
slpA10	P	N	H	A	T	T	P
slpE10	K	R	G/E	M	S	G	T

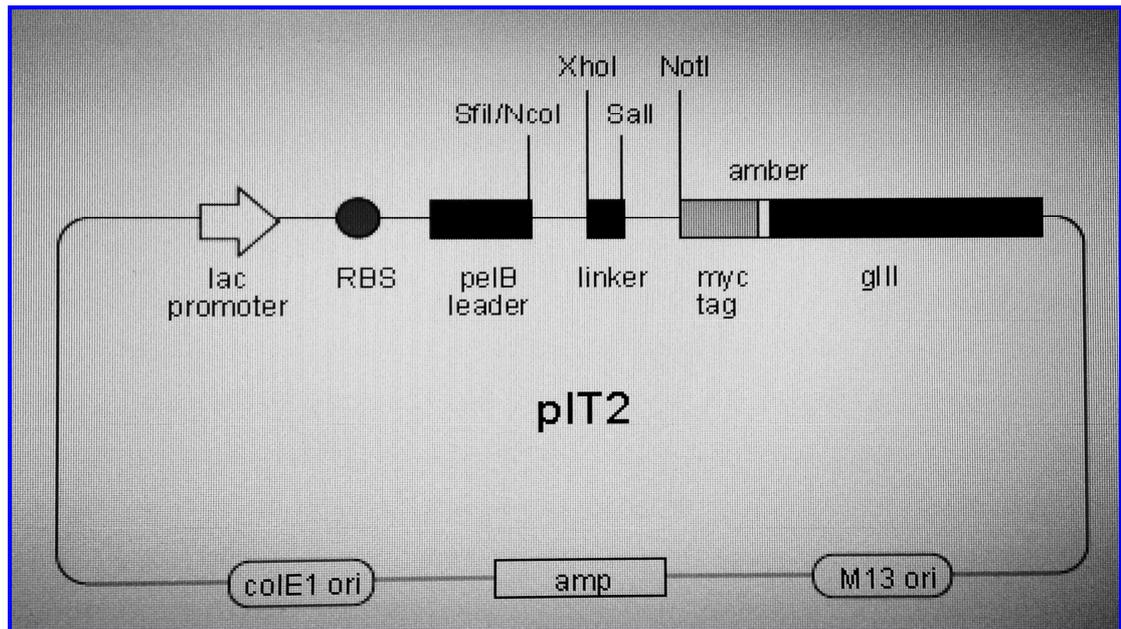


Figure 3. 65. Vector map of phagemid pIT2

ScFv is cloned between pelB leader and gIII protein of phage M13.

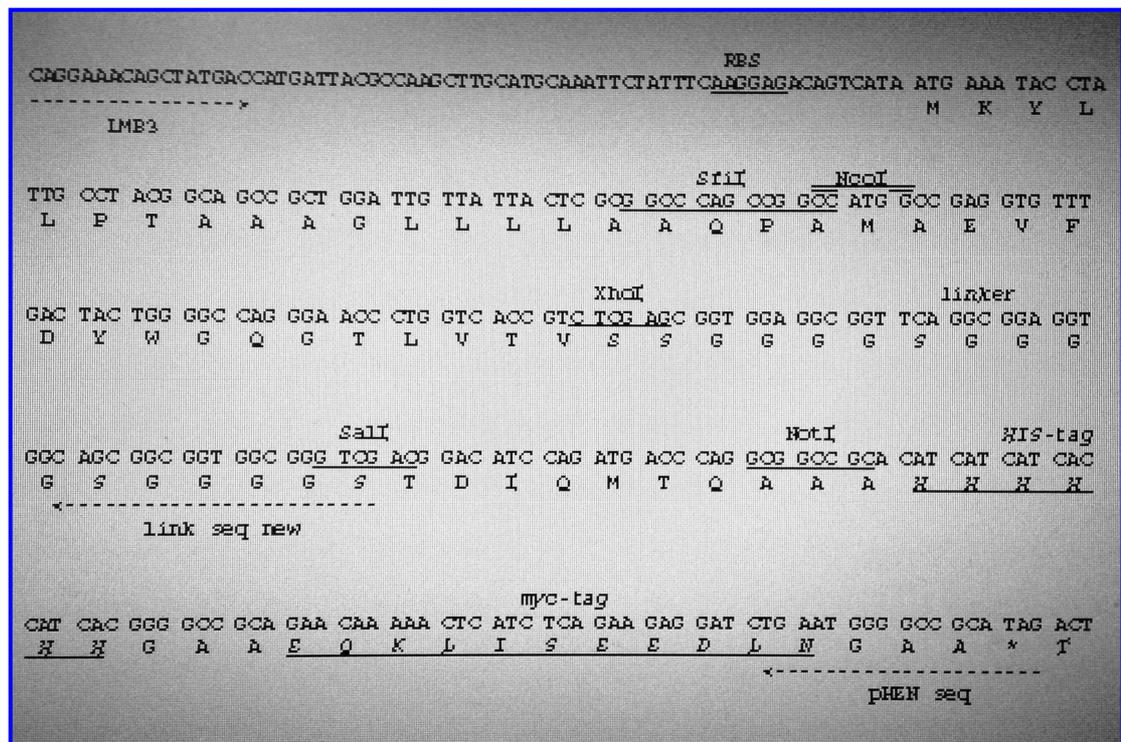


Figure 3. 66. Sequence region of restriction sites on phagemid pIT2.

VH sequence is inserted between *NcoI* and *XhoI* sites and VL lies between *SalI* and *NotI*. Each scFv includes a his tag and c-myc-tag before the TAG stop codon. The full length scFv can be excised by digestion with *NcoI* and *NotI* sites while

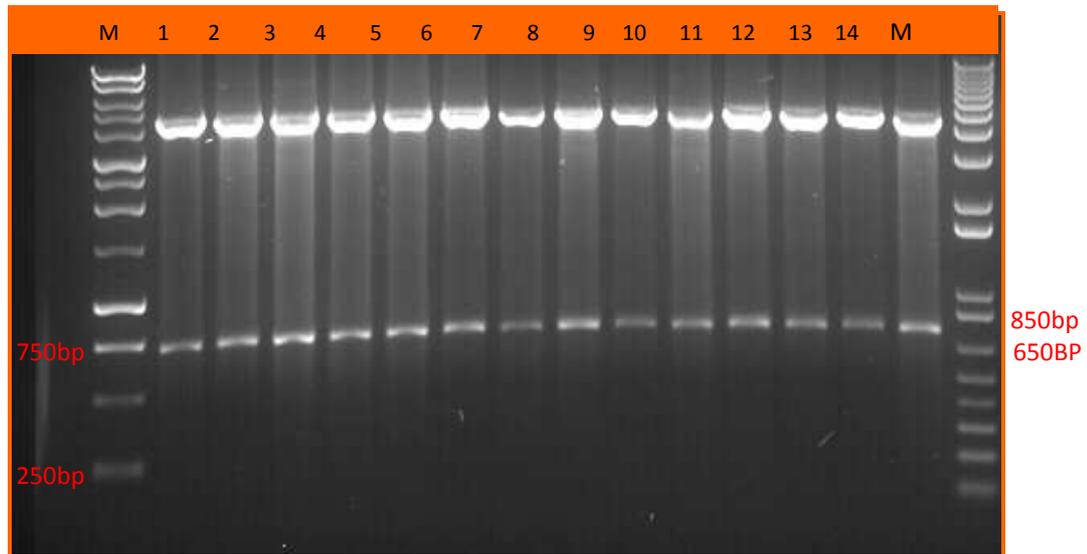


Figure 3. 67. Restriction analysis for the presence of full-length scFv inserts in clones from the Tomlinson library.

Plasmid DNA from each clone was digested with *NcoI* and *NotI*. Release of a full-length scFv insert is expected to result in a fragment of 717 bp.

From the left: M, 1kb DNA ladder (Promega); Lane 1, N8B4; lane 2, N8B5; lane 3, N8A6; lane 4, N6B5; lane 5, N6F1; lane 6, LF10; lane 7, LH1; lane 8, LG7; lane 9, LG1; lane 10, LA6; lane11, LD4; lane 12, slpA10; lane 13, slpE10; lane 14, slpA9; M, 1kb plus marker (Invitrogen).

Clones designated N6 were isolated by panning against FliD, N8 are directed against FliC, L indicates binding to LMW, and clones designated slp were isolated against slpA.

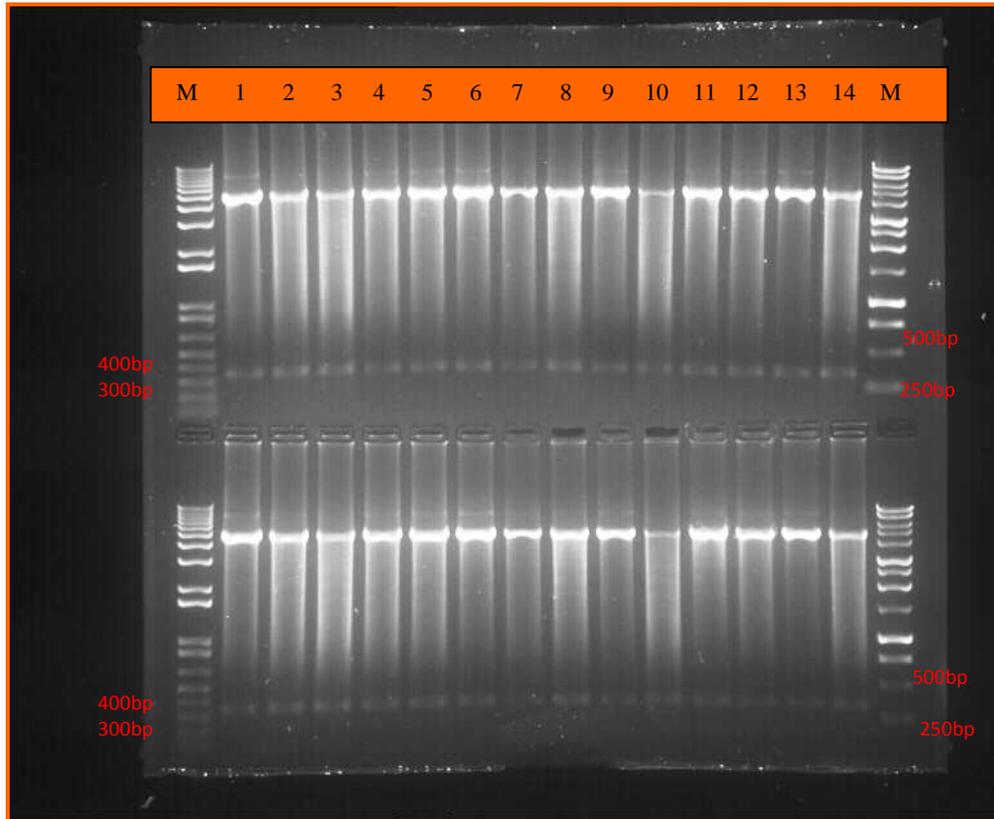


Figure 3. 68. Restriction analysis of selected clones from Tomlinson library.

Upper gel: Plasmid of each clone was digested with NcoI and XhoI . The expected appearing fragment is 340 bp. Lower gel: digestion with Sall and NotI with expecting a 325 bp fragment size. From left: 1kb puls marker (Invitrogen); Lane 1, N8B4; lane 2, N8B5; lane 3, N8A6; lane 4, N6B5; lane 5, N6F1; lane 6, LF10; lane 7, LH1; lane 8, LG7; lane 9, LG1; lane 10, LA6; lane11, LD4; lane 12, slpA10; lane13, slpE10; lane 14, slpA9; 1kb DNA ladder (Promega). N6=fliD, N8=fliC, L=LMW AND slp=slpA.

3.3.6 Antibody purification

Recombinant antibodies from the Tomlinson libraries are expressed as monovalent scFv proteins with histidine and c-myc tags at carboxy-termini (Figure 3.66). The histidine tag can be used to purify the soluble fragments through nickel affinity chromatography and as detailed earlier, immunodetection of the presence of the protein is possible using reagents directed towards the c-myc component.

A large culture (200-500 ml) of each *E. coli* HB2151 clone carrying a specific scFv was induced with IPTG. Supernatants from overnight cultures were concentrated with 80% ammonium sulphate or ultrafiltration concentrators with a 10 kDa cut-off in order to reduce the volume of sample to be processed. Concentrated fractions or dialysed samples from ammonium sulphate precipitation were passed through an equilibrated nickel chelating column allowing the scFv protein to bind. To elute the attached proteins, 200mM imidazole was pumped through the column and at least 5 fractions were collected. Fractions were initially checked *via* dot blotting to assess for the presence of protein with the c-myc tag. Those fractions that showed positive signals were pooled.

Figures 3.69 and 3.70 show dot blots of fractions from purification of scFvs against FliC, FliD, LMW and SlpA. Based on the blot analysis, and the green colorimetric signal, scFv was recovered from each purification in a variable number of fractions. Pooled fractions from each purification were sampled and analysed by SDS-PAGE. Figures 3.71 and 3.72 show the presence of a band of 30kDa, the expected molecular weight for a scFv protein. The concentration of the scFvs and their purity were variable. For example, the anti-FliD scFv N6C9 (lane 4 in Figure 3.71) was detectable but recovered at low concentration when compared to other anti-FliC and anti-FliD antibodies shown in Figure 3.71. scFv LA6 (lane 4, Figure 3.72) was also recovered at low concentration. The scFvs against FliC (lanes 5 to 9, Figure 3.71) were generally recovered with good purity, accepting that in lane 9

(scFv N8B4) trace impurities are evident but the overall concentration of scFv appears higher than in other samples. In contrast, anti-FltD antibodies were typically recovered with contaminating proteins of variable molecular weight (eg N6B5 [lane 1, Figure 3.71] *versus* N6F1 [lane 3 in the same Figure]) and for N6A5 (lane 2), the contaminant was present at roughly equal concentration to the scFv. Although greater consistency of purity was achieved for scFvs against LMW and SlpA (Figure 3.72), the antibodies were recovered at a range of concentrations (eg LA6 [lane 4] was present at very low concentration whereas LG7 [lane 2] was more abundant). Irrespective of outcome, pooled fractions were concentrated using an ultrafiltration concentrator with a 10 kDa cut-off and stored at -20°C pending further analysis.

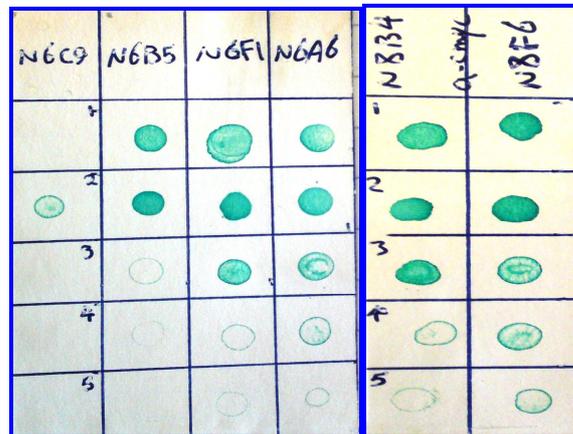


Figure 3. 69. Dot blot analysis from purification of anti FliC (N6) and anti FliD (N8) scFvs. Spotted samples on the Hyband C membrane were from different fractions of purification scFv through the Ni-chromatography. The presence of antibodies were detected with anti-c-myc reagent. The green signal was developed by adding anti-rabbit-HRP and TMB.

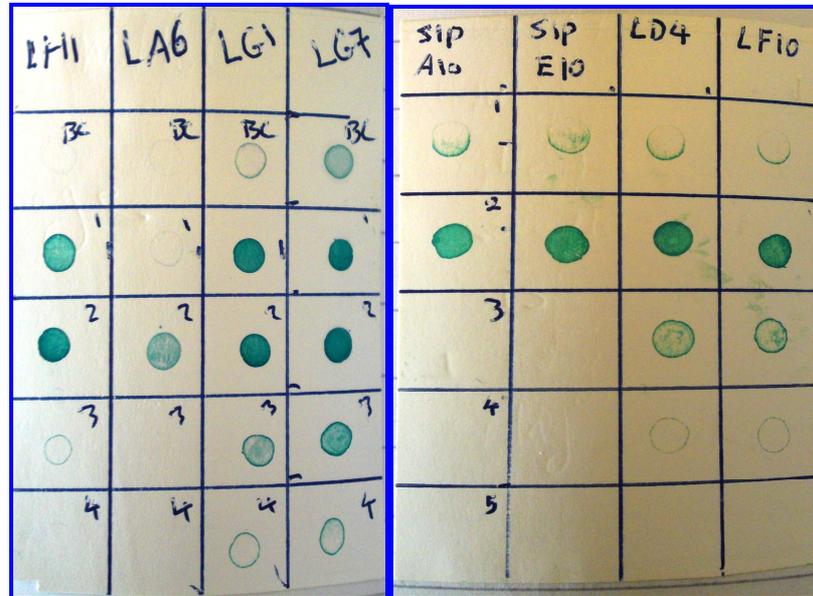


Figure 3. 70. Dot blot analysis from purification of anti LMW and anti SlpA scFvs. Spotted samples on the Hyband C membrane were from different fractions of purification scFv through the Ni-chromatography. Presence of antibodies was detected with anti-c-myc reagent after blocking with 2%MPBS. The green signal was developed by adding anti-rabbit-HRP and TMB.

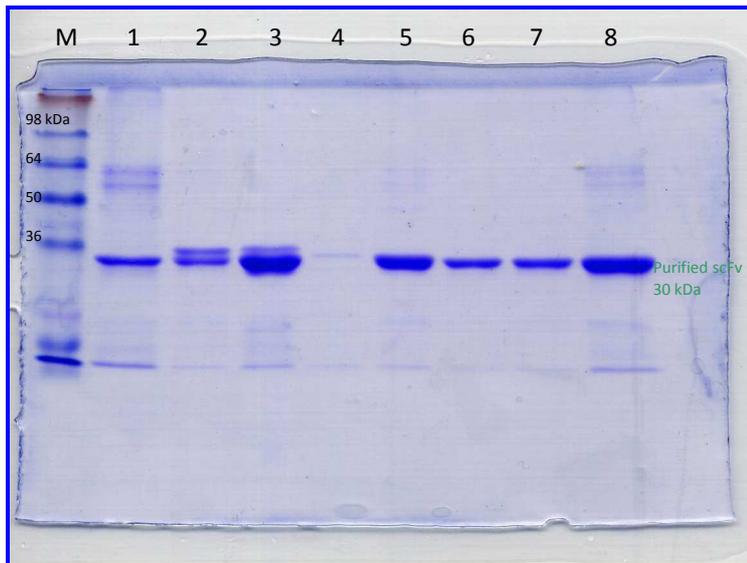


Figure 3. 71. SDS-PAGE analysis of purified scFvs against recombinant FliC and FliD.

Columns were loaded with sample after adjusting pH and buffer conditions and washed with binding buffer initially without imidazole and then with buffer containing a low concentration of imidazole (40 mM). The scFv protein was eluted with 200 imidazole in binding buffer.

From the left: Marker (SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen)); lane 1, N6B5; lane 2, N6A6; lane 3, N6F1; lane 4, N6C9; lane 5, N8A6; lane 6, N8B5; lane 7, N8F6; lane 8, N8B4.

scFvs designated N6 are directed against FliD, N8 indicates anti-FliC antibodies.

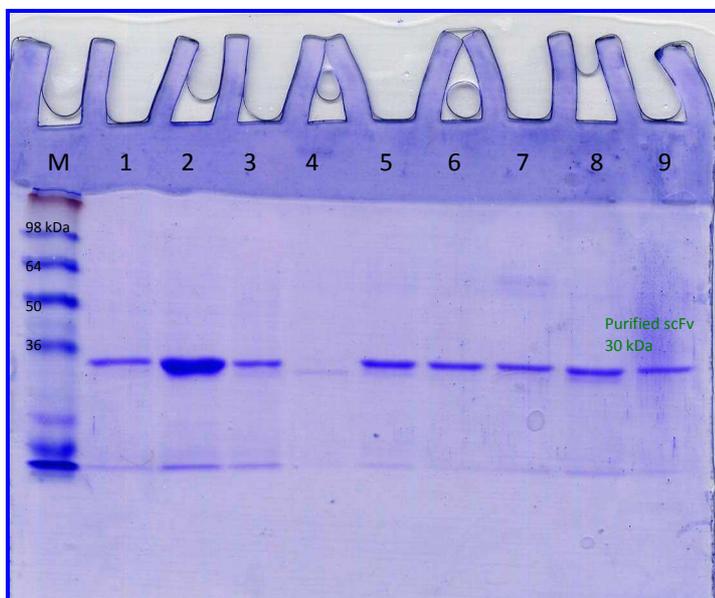


Figure 3. 72. SDS-PAGE analysis of pooled purified fractions of scFvs against recombinant LMW and extracted SlpA. Columns were loaded with sample after adjusting pH and buffer conditions and washed with binding buffer initially without imidazole and then with buffer containing a low concentration of imidazole (40 mM). The scFv protein was eluted with 200 mM imidazole in binding buffer. From left: Marker (SeeBlue Plus2 Pre-Stained molecular weight markers (invitrogen), lane 1, LF10; lane 2, LG7; lane 3, LD4; lane 4, LA6; lane 5, LG1; lane 6, LH1; lane 7, slpE10; lane 8, slpA10; lane 9, slpA9. scFvs designated L are directed against LMW, slp indicates anti-slpA antibodies.

3.3.7 *Western blot analysis of selected scFvs against targets*

3.3.7.1 Immunoblot analysis with recombinant target proteins

Polyacrylamide gel electrophoresis and Western blotting were carried out to determine the specificity of the purified recombinant antibodies and to assess if they would recognise the clostridial target proteins after denaturation in SDS. Recombinant target proteins were run on SDS gels and then blotted to Hybond membrane which was then divided and individual strips were probed with scFvs against the target. The binding of scFvs was detected with anti c-myc antibody and anti rabbit-HRP as the developing reagent. In preliminary experiments, a mixture of three scFvs directed against a particular target was used in Western blot to reduce the number of blots required to screen all scFvs. Having established in these experiments that the scFvs would work on Western blots, single scFvs were then used. Figures 3.73 – 3.80 show the results of these experiments.

The first general finding that is apparent through all Figures is that the scFvs were capable of recognising the recombinant clostridial proteins despite their treatment with SDS sample buffer and heating in preparation for SDS-PAGE. As regards the properties of groups of scFvs, signal strength in the blots was generally high with the exception of reactions against FliC (Figure 3.77). This was unlikely to be due to limitations of target on the blot since the FliC protein was successfully expressed and purified (Figure 3.29, lane 18). In many cases, cross-reaction or non-specific binding to other proteins on the blots was minimal but in Figure 3.76, some differences can be observed between anti-FliD scFvs N6A4 and N6B5 (left and central panels; little binding to other proteins on the blot), and N6C9 which appears to bind to other species (right panel). In contrast, the anti-Cwp66 antibodies N11A5, N11C1 and N11F4 appear more consistent in their binding to a contaminating protein of about 64 kDa (Figure 3.80).

3.3.7.2 Immunoblot analysis with native *C. difficile* proteins

With the exception of the anti-SlpA antibodies, all scFvs isolated from the Tomlinson libraries used recombinant clostridial proteins as targets and hence their ability to recognise the native proteins from *C. difficile* remained unknown to this point in the project. To address this issue, extracts from overnight cultures of *C. difficile* were prepared, run in SDS-PAGE and transferred to membranes so that the recognition of native proteins could be assessed *via* Western blotting. As an alternative approach for scFs directed against components of the clostridial surface layer, native electrophoresis was also carried out to determine if the behaviour of these antibodies changed when SDS was omitted from the electrophoretic separation.

Figure 3.81 shows the recognition by a range of scFvs of SlpA extracted from the surface of *C. difficile* 630. Note that scFvs with the “L” designation were originally isolated by panning with recombinant LMW protein whereas those with the “Slp” were extracted from the Tomlinson libraries using clostridial extracts. Reference to the middle row in Figure 3.81 illustrates that the dominant components of the bacterial extracts are two proteins, one of them under 50 kDa and one of a lower molecular weight (around 36 kDa). In the middle row, it can be seen that scFv LG1 (left panel) binds to the heavier of these proteins, whereas LF10 (middle panel) and LG7 (right panel) recognise the lower molecular weight component. For LF10 and LG7, faint signal can be observed towards the 50 kDa molecular weight marker confirming this pattern. Other scFvs isolated by panning on the LMW protein (LD4, LH1, LA6; top row of the Figure) were consistent in their recognition of the higher molecular weight protein. Two scFvs isolated from panning on SlpA extracts (SlpA9 and SlpE10) recognised the lighter species (bottom row of the Figure) whilst a third bound to the 50 kDa form. In summary, all the anti-SlpA scFvs recognised native target protein from *C. difficile* 630 in conventional Western blotting but whereas LA6, LD4, LG1, LH1 and SlpA10 could bind to the high molecular weight component, LF10, LG7, SlpA9 and SlpE10 recognised the smaller protein.

Since the sequences of SlpA proteins possess conserved and variable regions, all nine scFvs against SlpA were tested in Western blotting using extracts from *C. difficile* 630 and two other strains, R20291 and M120 (Figure 3.82). None of the scFvs were able to bind to SlpA from *C. difficile* R20291 (central lane in each panel) but all showed the capacity to bind to SlpA components from strains 630 and M120 and the characteristics described earlier – the ability to recognise high and low molecular weight components of SlpA – was also evident. Recombinant antibodies LF10, LG7, SlpA9 and SlpE10 could all bind to the low molecular weight protein from 630 and M120. The remaining scFvs, (LA6, LD4, LG1, LH1 and SlpA10) recognised the high molecular weight component of SlpA from strain 630 but also the smaller protein from M120.

Thus, the binding of scFvs to SlpA from *C. difficile* 630 was consistent with data presented earlier (Figure 3.81) but in contrast, all scFvs bound to the low molecular weight component of SlpA from strain M120. In some cases, this recognition resulted in strong signals on Western blots (scFvs LD4 and LF10; Figure 3.82, top row central and right hand panels respectively, note the right lane in each case) whereas other scFvs reacted less strongly with material from M120 *versus* 630 (LG1; middle row, left hand panel). In some instances, some recognition of the higher molecular weight protein in SlpA from *C. difficile* M120 was also evident (LG7 [middle row, central panel], SlpA9 [bottom row, left panel], SlpA10 [bottom row, right panel]) but signal strength was weak in these cases.

Further analysis was pursued by separating proteins by native electrophoresis and Western blotting. Using this approach the separation of high and low molecular weight constituents of the SlpA of *C. difficile* 630 was less apparent and the ability of scFvs LF10 and LG7 to react with the lower molecular weight component of SlpA described earlier was lost (Figure 3.83 top row, right panel; middle row, central panel). Other scFvs shown to be reactive with this protein (SlpA9, SlpE10; Figure 3.83 lower row left and right panels respectively) retained some reactivity with a low molecular weight constituent but binding to

the main complex appeared dominant. Also, scFv LH1, shown in previous experiments to recognise the higher molecular weight protein from SlpA, showed some reaction with a lighter species (Figure 3.82 middle row right panel) so non-specific binding might underlie the result. Native electrophoresis again showed that none of the panel of scFvs were able to bind to extracts from *C. difficile* R20291. Interaction with a protein from M120 of comparable molecular weight to the smaller protein from SlpA was retained. This was evident for all scFvs tested (Figure 3.83).

The scFvs which were isolated against recombinant FliC and FliD were also tested against native forms of the proteins using crude flagellar extracts from *C. difficile* and Western blotting. For the scFvs against FliD, no signal was detected in Western blotting. Given the strong reaction of scFvs N6A4, N6B5 and N6C9 with recombinant FliD (Figure 3.76), the result with bacterial extracts most probably arose because of the low amounts of FliD in the flagellar preparations. In contrast with Figure 3.77, Figure 3.83 shows the scFvs against recombinant FliC could bind to native proteins with molecular weights between about 28 and 38 kDa that were present in the flagellar extracts from *C. difficile* 630. All three scFvs had an ability to bind to these proteins; the complexity of the pattern of signals was unexpected and reaction with a protein so close to the 38 kDa molecular weight markers was unexpected since FliC from *C. difficile* 630 is known to have a molecular weight of around 33 kDa owing to glycosylation of the protein (predicted molecular weight of the translation product is 30.7 kDa). The heaviest band in the profile was notable for its similarity of size to the LMW component of SlpA. Binding to a protein in flagellar extracts from strain R20291 was weaker but still detectable.

Given the crude nature of the surface and flagellar extracts, the extent of cross-contamination and potential cross-reaction was explored using Western blotting with an anti-FliD scFv (N8B5) and LG7 which was isolated against a recombinant preparation of the LMW component of SlpA. Data is shown in Figure 3.85. Panels 1 and 2 carry flagellar extracts of *C. difficile* 630 and the blots were

probed with N8B5 and LG7 respectively. The target in panels 3 and 4 were surface extracts; again, these were probed with N8B5 (panel 3) and LG7 (panel 4). The alignment of molecular weight makers in the Figure is imperfect but it appears from comparison of panels 1 and 2 that N8B5 is reactive with a component of the flagellar extract of similar molecular weight to the contaminating LMW component of SlpA present in these extracts. Cross reaction between N8B5 and LMW seems unlikely from the patterns of reaction shown in panels 3 and 4.

Figure 3.86 presents an analysis of the properties of recombinant antibodies against CspA. The lysates of three strains (630, M120 and R20291) were used as targets from the blots. The Western blots showed that binding was only evident from extracts of M120 and in each case, antibody attachment appeared to take place to LMW of SlpA of M120 or a protein of similar molecular weight. The predicted molecular weight of native CspA is 7 kDa. In Figure 3.87 illustrates the reaction of scFvs against GroEL with proteins in extracts from the same strains of *C. difficile*. A signal at 60 kDa suggests specific reaction with GroEL in extracts from all strains (left panel) with scFv antibody N5E10. Other weak signals were developed around 50 kDa in 630 extracts for all three scFvs but the strongest signals are at 36 kDa, against suggestive of a reaction with the LMW component of SlpA from strain M120. Therefore based on Figure 3.87, three independent scFvs against GroEL appear able to bind to LMW of SlpA from M120. In Figure 3.88 the scFvs against recombinant SortaseB were tested in Western blots. A band of around 27 kDa was expected in this experiment. As the Figure shows, the three anti-Sortase scFvs also bound strongly to the LMW from M120 strain, although weaker signals from a 50kDa protein can be also be seen.

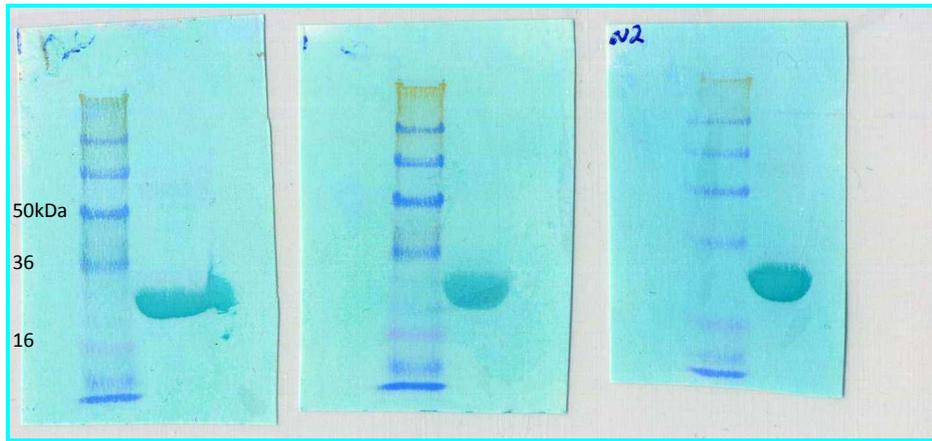


Figure 3. 73. Western blot analysis of CspA probed with scFvs

Recombinant CspA (24.96 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N2A9; middle panel, N2E7; right panel N2G12.

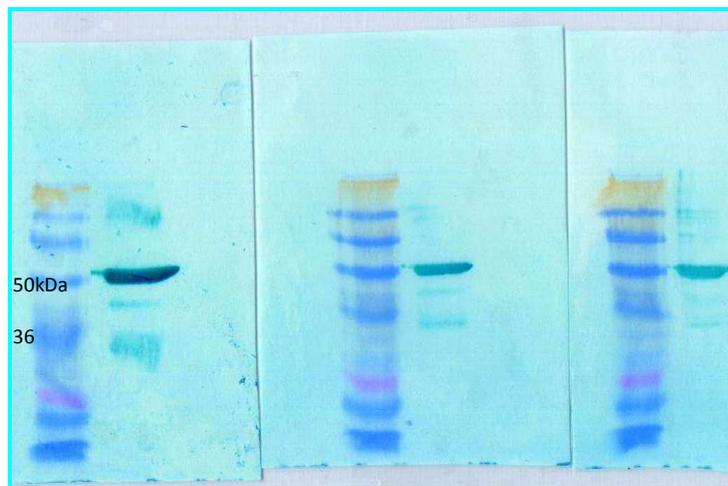


Figure 3. 74. Western blot analysis of N terminal of Cwp66 probed with scFvs

Recombinant N-terminal of cwp66 (47.07 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N3A1; middle panel, N3F4; right panel N3H12.

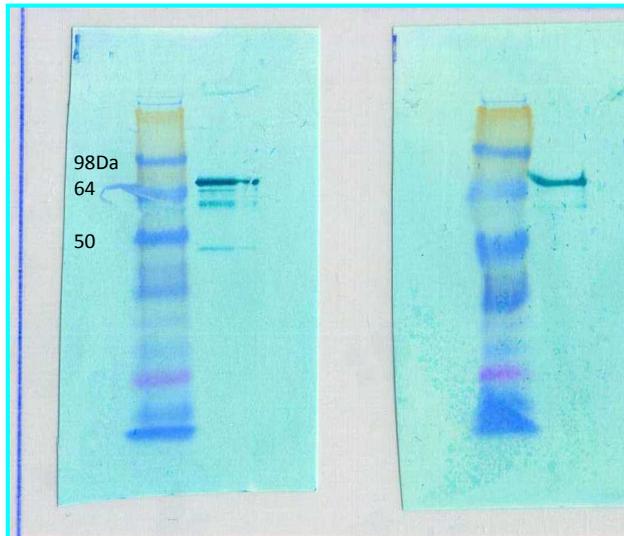


Figure 3. 75. Western blot analysis of GroEL probed with scFvs

Recombinant groEL (72.2 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N5E10; right panel; N5E9.

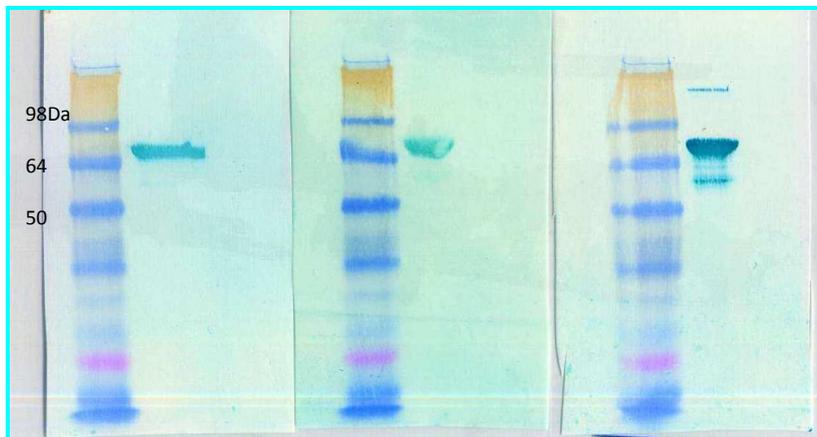


Figure 3. 76. Western blot analysis of Flid probed with scFvs

Recombinant flid (75.78 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N6A4; middle panel, N6B5; right panel N6C9.

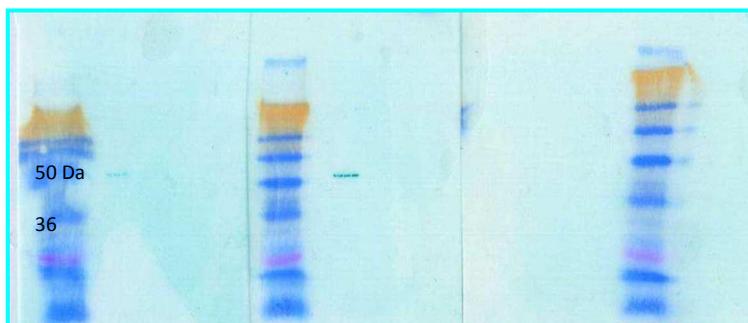


Figure 3. 77. Western blot analysis of FliC probed with scFvs

Recombinant fliC (48.04 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N8F6; middle panel, N8B4; right panel N8A1.

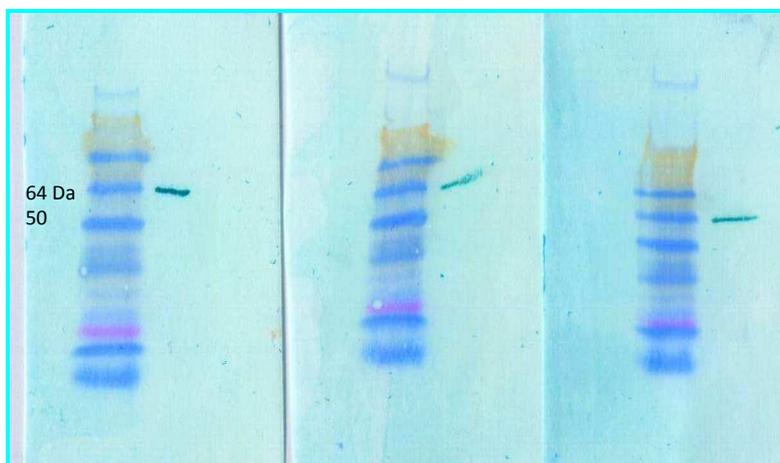


Figure 3. 78. Western blot analysis of N-terminal of Cwp66 probed with scFvs

Recombinant N-terminal of cwp66 (53.94) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N3A1; middle panel, N3F4; right panel N3H2.

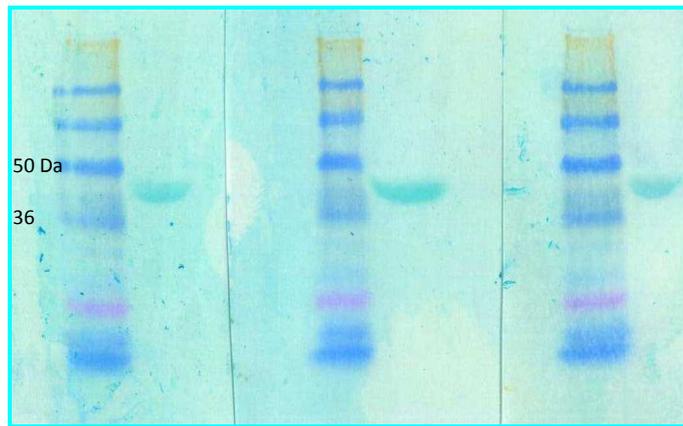


Figure 3. 79. Western blot analysis of Sortase B probed with scFvs

Recombinant sortase b (43.38 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N10D11; middle panel, N10D10; right panel N10B8.

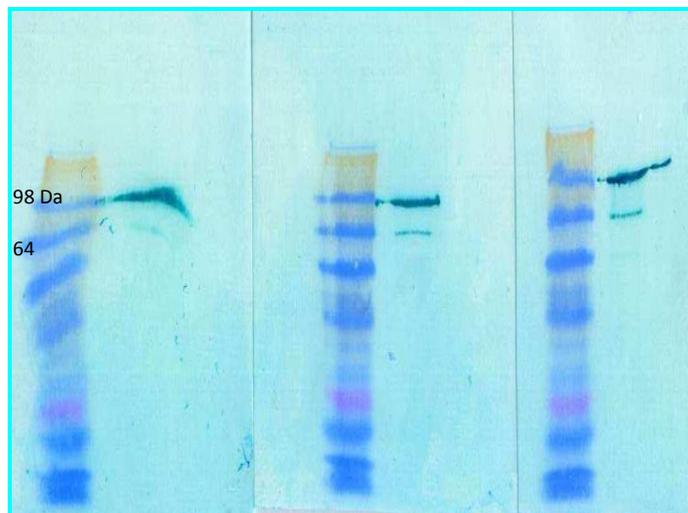


Figure 3. 80. Western blot analysis of Cwp66 probed with scFvs

Recombinant cwp66 (83.31 kDa) was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N11A5; middle panel, N11C1; right panel N11F4.

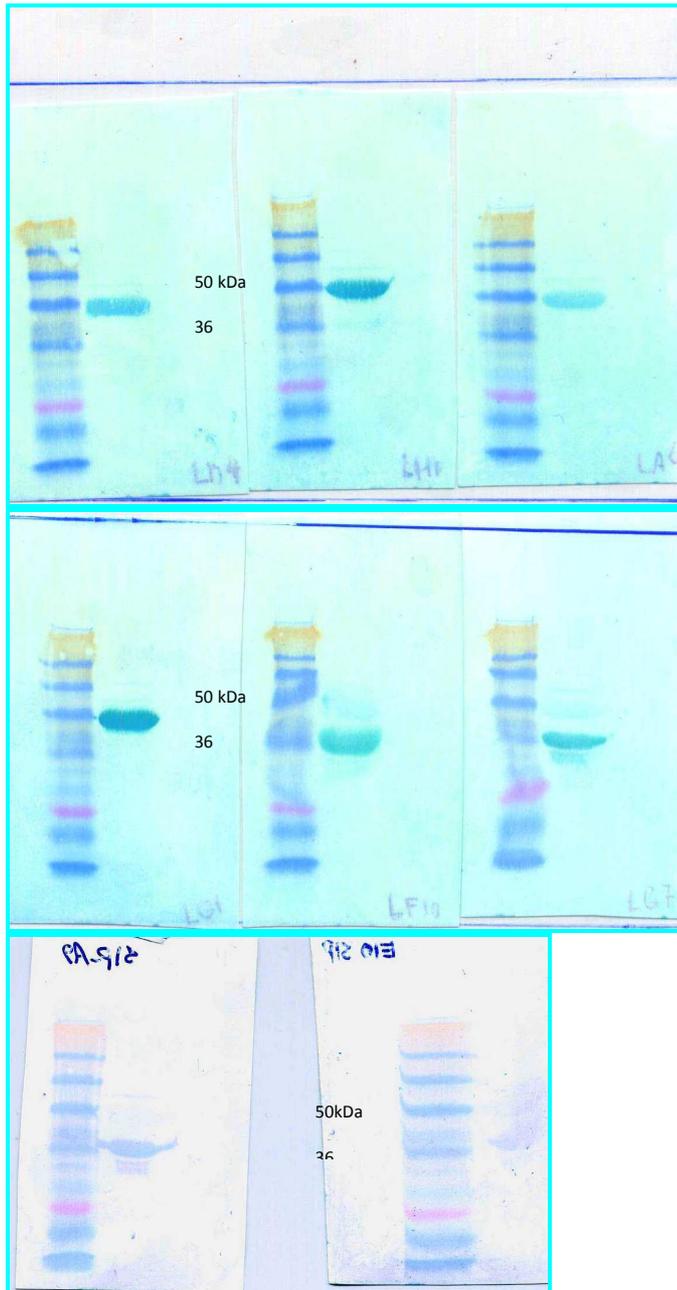


Figure 3. 81. Western blot analysis of SlpA extracted from *C. difficile* 630 probed with scFvs

Extracted SlpA from *C. difficile* 630 was separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen).

Upper row: left panel, LD4; middle panel, LH1; right panel, LA6.

Middle row: left panel, LG1; middle panel, LF10; right panel, LG7.

Bottom row: left panel, SlpA9; right panel, SlpE10.

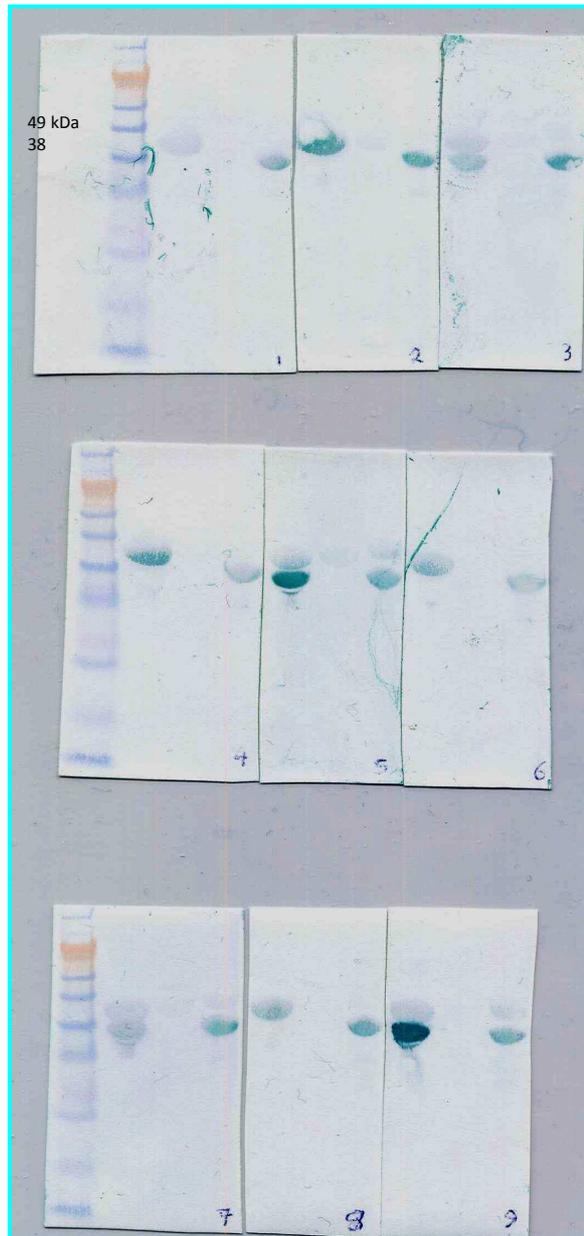


Figure 3. 82. Western blot analysis of SlpA extracted from *C. difficile* 630, R20291 and M120 probed with scFvs

Extracted SlpA from *C. difficile* 630, R20291 and M120 were separated by SDS-PAGE and transferred to Hybond membrane that were then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen).

Upper row: left panel, LA6; middle panel, LD4; right panel, LF10.

Middle row: left panel, LG1; middle panel, LG7; right panel, LH1.

Bottom row: left panel, SlpA9; right panel, SlpE10.

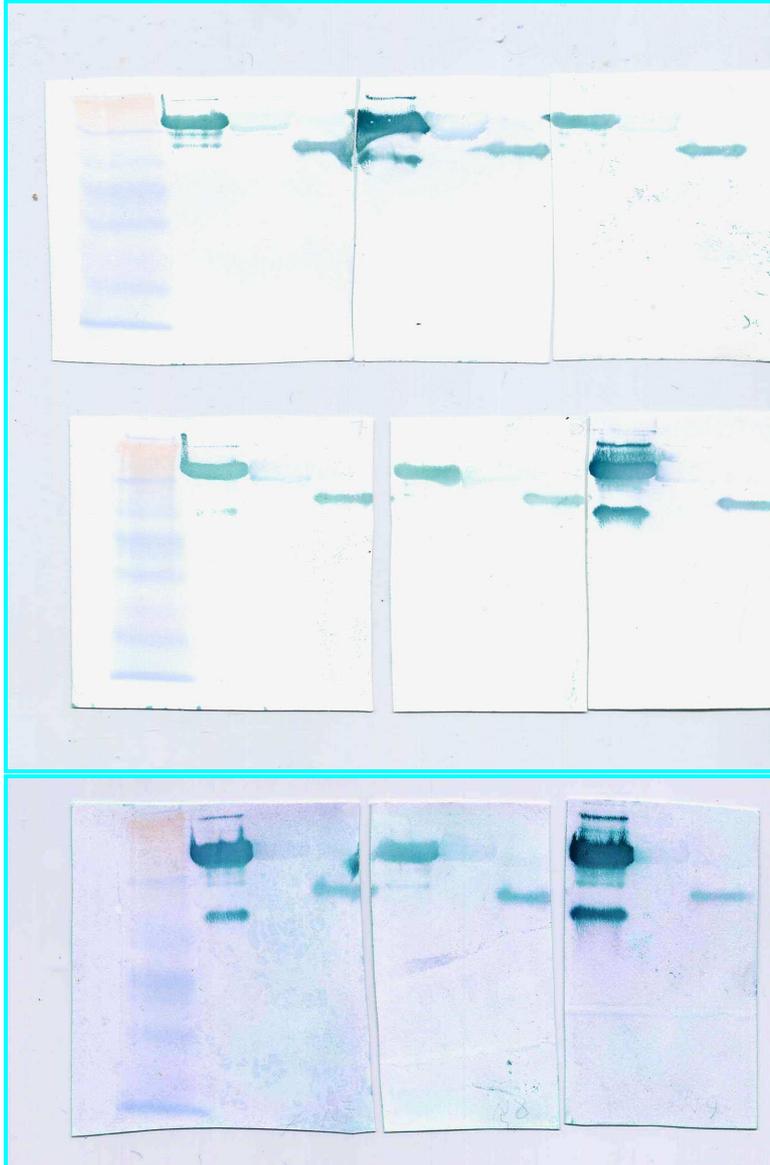


Figure 3. 83. Western blot analysis of SlpA extracted from *C. difficile* 630, R20291 and M120 probed with scFvs

Extracted SlpA from *C. difficile* 630, R20291 and M120 were separated by 8% native electrophoresis and transferred to Hybond membrane that were then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen).

Upper row: left panel, LA6; middle panel, LD4; right panel, LF10.

Middle row: left panel, LG1; middle panel, LG7; right panel, LH1.

Bottom row: left panel, SlpA9; middle panel, SlpA10; right panel, slpE10.

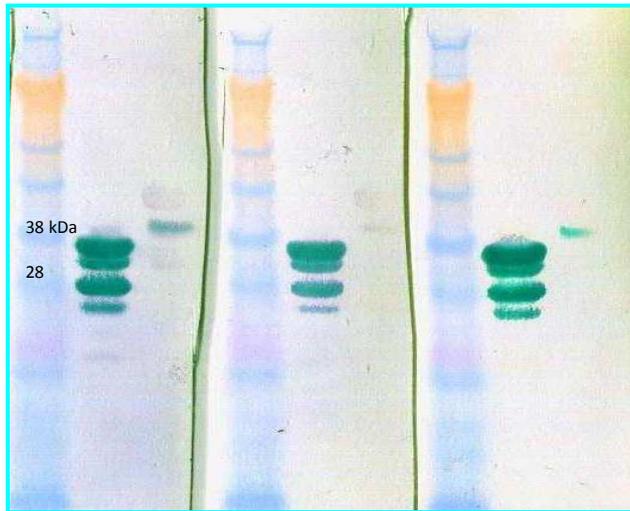


Figure 3. 84. Western blot analysis of flagellum preparation from *C. difficile* 630 and R20291 probed with scFvs.

Flagellum preparation from *C. difficile* 630 and R20291 were separated by SDS electrophoresis and transferred to Hybond membrane that were then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen).

Left panel, N8A6; middle panel, N8B4; right panel, N8B5.

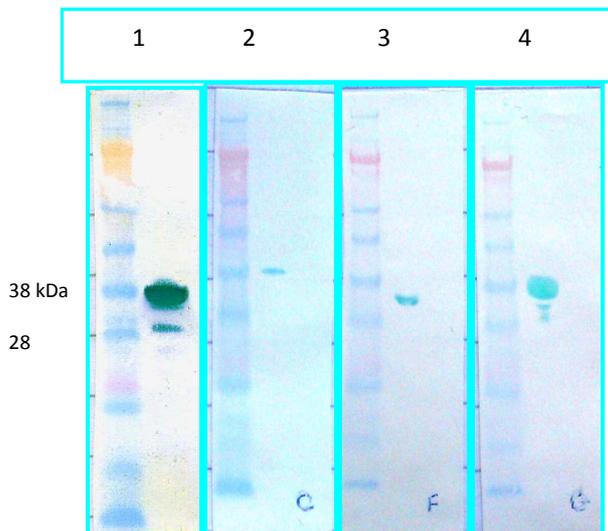


Figure 3. 85. Western blot analysis of FliC and LMW extracted from *C. difficile* 630 probed with scFvs.

Extracted fliC and SlpA from *C. difficile* 630 were separated by SDS-PAGE and transferred to Hybond membrane that were then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight markers (Invitrogen).panel 1, N8A6; panel 2, LG7; panel 3, N8B5; panel 4, LG7.

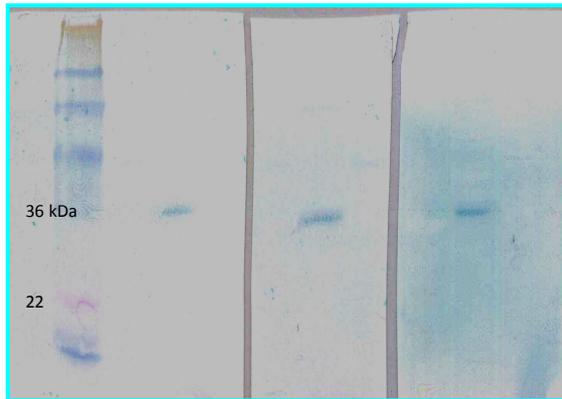


Figure 3. 86. Western blot analysis of CspA from *C.difficile*, 630, M120 and R20291 probed with scFvs.

Bacterial lysate of three strains were separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N2A9; middle panel, N2E7; right panel N2G12.

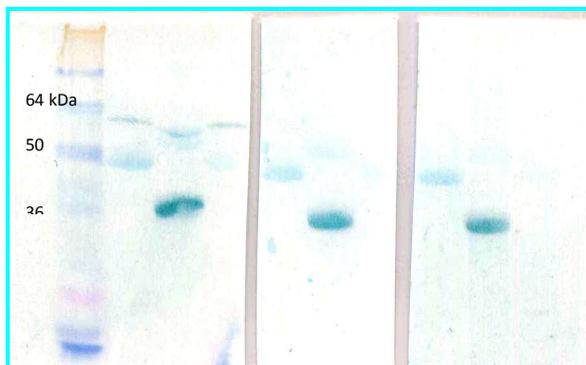


Figure 3. 87. Western blot analysis of CspA from *C.difficile*, 630, M120 and R20291 probed with scFvs.

Shocked bacterial lysate (42 °C) of three strains were separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N5E10; middle panel, N5A1; right panel N5E9.



Figure 3. 88. Western blot analysis of CspA from *C.difficile*, 630, M120 and R20291 probed with scFvs.

Bacterial lysate of three strains were separated by SDS-PAGE and transferred to Hybond membrane that was then probed with an individual scFv. Anti c-myc and an anti rabbit-HRP conjugate was used to detect the binding of scFv.

In each panel, the left hand lane shows the migration of SeeBlue Plus2 pre-stained molecular weight standards (Invitrogen). The sizes of relevant molecular weight markers are indicated. scFvs used for the analysis were as follows:

Left panel: N10B8; middle panel, N10D11; right panel N10D10.

3.3.8 Analysis of the effect of antibodies on bacterial motility

To determine the biological effects of scFvs against FliC and FliD, bacterial motility tests were carried out in the presence of these antibodies. To each counted inoculum of overnight culture of *C. difficile* 630 (10^5 CFU per sample), scFv was added and after incubating for one hour in anaerobic cabinet, 10 μ l of the mixture was stabbed into 0.2% BHI agar. LA6 was used as a control for a scFv which was not reactive against flagellar proteins. The stabs prepared without antibody or inocula mixed with LA6 showed good evidence of the bacterial motility. Of the nine scFvs against FliC and the nine anti-FliD scFvs, four anti-FliD antibodies and four anti-FliC scFvs could inhibit bacterial motility.

Based on the results shown in Table 3.21, in total five antibodies were chosen for the second part of the inhibition test. Figure 3.89 shows the effects of serial dilution of the different scFvs on motility of bacteria. Each row shows 3 tubes each stabbed with 10^5 CFU following incubation of the bacterial cells with undiluted scFv and 1/10 and 1/100 dilutions of the antibodies. No motility could be seen by at any of the three dilutions after 48 hours incubation. In the last row of the Figure, tubes show the diffuse pattern of growth observed with a culture of *Clostridium difficile* 630 (left), the same strain after incubation with the control scFv LG6, and a stab prepared with the non-motile strain M120. Based on comparisons, strong evidence was obtained that the scFvs against FliC and FliD could inhibit bacterial motility.

Table 3. 21. Comparing of the effect of different scFvs on mobility of *C.difficile* 630

scFvs AGAINST FliD	24H	48H	scFvs AGAINST FliC	24H	48H
N6A4	SM	SM	N8A1	SM	SM
N6A6	SM	SM	N8A6	NM	NM
N6A8	SM	SM	N8B4	NM	NM
N6B5	NM	NM	N8B5	NM	NM
N6C9	NM	NM	N8F1	SM	SM
N6D4	NM	NM	N8F6	NM	NM
N6D12	SM	SM	N8H2	SM	SM
N6F1	NM	NM	N8H5	SM	SM
N6H11	SM	SM	N8H11	SM	SM

*SM=some motility, NM=non motile

	N6F1
	N6B5
	N8A6
	N8B4
	N8B5
	From left: <i>C. difficile</i> 630, LA6+bacteria and <i>C. difficile</i> M120

Figure 3. 89. Effect of recombinant antibodies on the motility of *C. difficile*

3.3.9 *Indirect immunofluorescence and electron microscopy*

In the final phase of the analysis, the scFvs were tested in immunofluorescence microscopy and immuno gold electron microscopy to assess if the surface location and distribution of target proteins could be determined by imaging methods. Experimental approaches included the use of Protein A-FITC to detect directly the presence of scFv at the bacterial surface or the use of anti c-myc and anti-rabbit conjugated reagents.

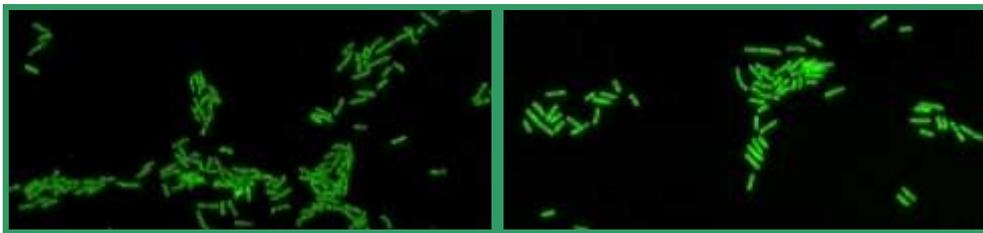
In immunofluorescence, most of the antibodies were found to be reactive with *C. difficile* M120, again supporting the notion that the LMW for SlpA of this strain was able to bind scFvs in a non-specific manner. N8F6 and N6F1 were carried out in this test for using *C. difficile* strains 630 and R20291. In both strains the selected scFvs generated strong signals at the bacterial surface (Figure 3.89). Of those antibodies able to bind to the LMW component of SlpA, scFvs LF10 and SlpA9 were tested in immunofluorescence with strains 630 and M120 strains. Both antibodies showed binding to the surface of bacterial cells (Figure 3.90). In these experiments, slides were also prepared without adding antibody (negative control) and mouse anti SlpA (positive control).

The resolution of these methods was insufficient to assess patterns of protein distribution. To address this, the antibodies were tested in immuno gold electron microscopy. Two labelling methods were tested: in the first, Protein A-gold was used in an attempt to detect the presence of scFv on sections; in the second, anti c-myc and anti rabbit-gold conjugates were used step-wise with the same objective. The first approach did not show any signs of gold particles on the grids while the second was more successful. In preparation for electron microscopy, bacteria were cultured in agar based medium and in liquid culture in an attempt to assess whether these conditions impacted upon the synthesis and assembly of proteins at the bacterial surface. Figures 3.91 and 3.92 show the results of attempts to label flagella and surface protein, respectively. The quality of the sections is good and while gold particles are clearly evident, their distribution does

not show evidence of the specific labelling of flagella (Figure 3.91) or high concentration of target proteins at the bacterial surface (Figure 3.92) as was evident in immunofluorescence.

A: N8F6-strain630

B: N6F1-strain 630



C: N8F6-strain R20291

D: N6F1-strain R20291

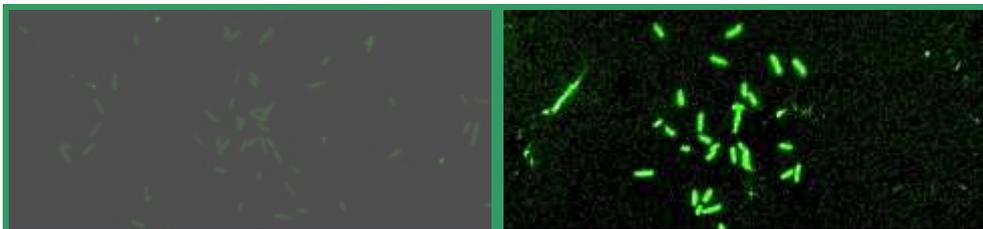
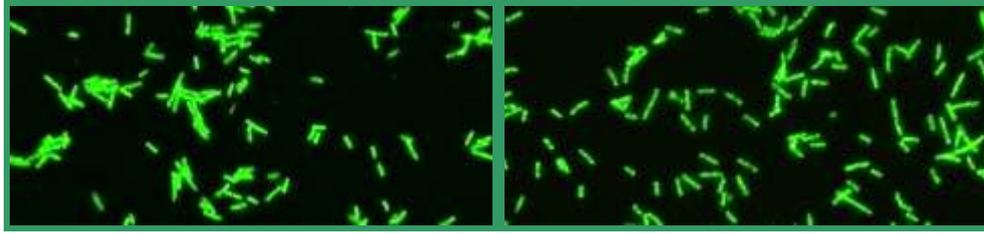


Figure 3. 90. Detection of flagellum of two strains (630 and R20291) in indirect immunofluorescence test. N6=fliC and N8=fliD. The overnight bacterial bodies were attached on the slide using cytopsin. The reagents were step by step added with interval PBS washing.

A: LF10-strain 630

B: slpA9-strain 630



A: LF10-strain M120

B: slpA9-strain M120

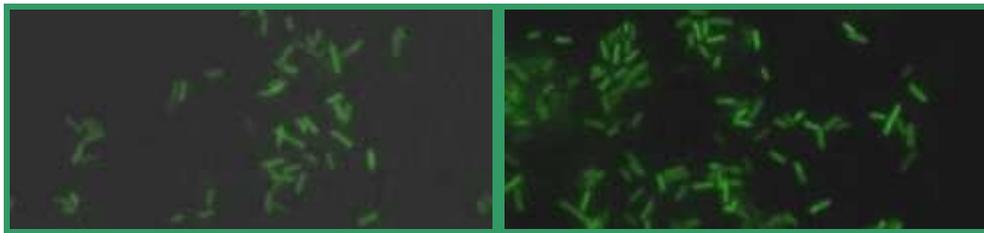
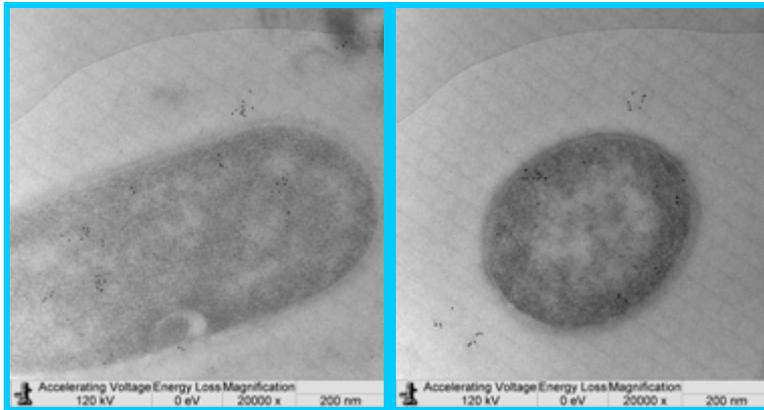


Figure 3. 91. Detection of surface layer protein on the surface of 630 and M120 strains in indirect immunofluorescence. Both checked antibodies were against LMW. The overnight bacterial bodies were attached on the slide using cytospin. The reagents were step by step added with interval PBS washing

A: N8F6



B: N6F1

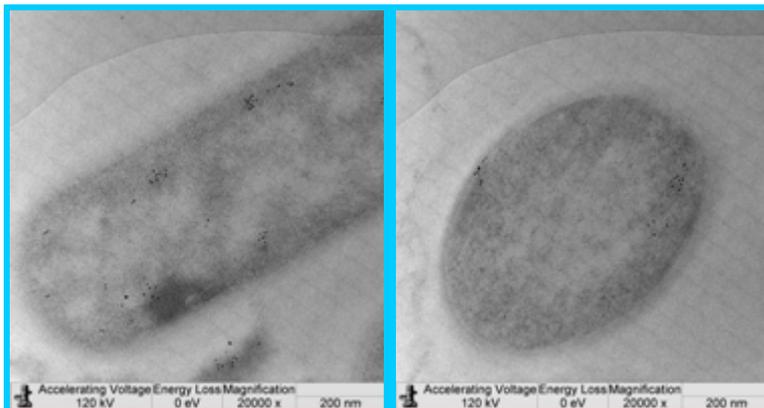
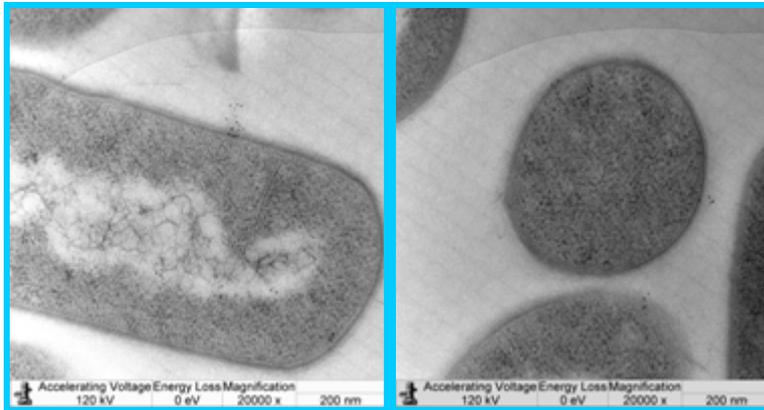


Figure 3. 92. Immunogold labelling of *C. difficile* 630 with recombinant antibodies against flagellum. Bacterial bodies were fixed and added on the grids from a plate agar based medium. The prepared grids were incubated with scFvs, anti c-myc and anti rabbit-gold respectively before examining in TEM. A) N6=fliC and B) N8=fliD

A: LF10



B: slpA9

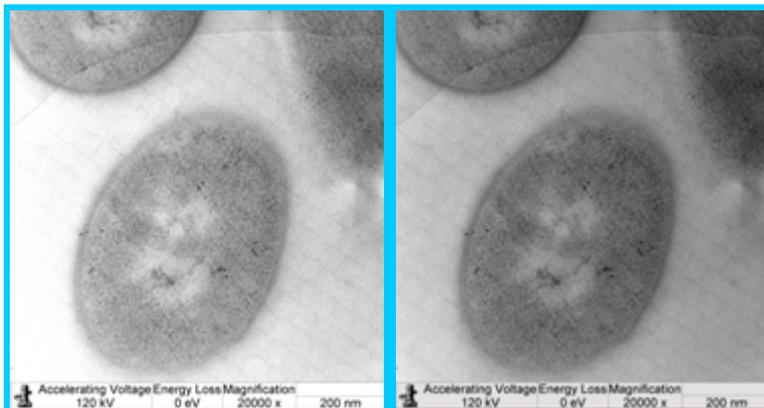


Figure 3. 93. Immunogold labelling of *C. difficile* 630 with recombinant antibodies against surface layer protein. Bacterial bodies were fixed and added on the grids from a liquid medium. The prepared grids were incubated with scFvs, anti c-myc and anti rabbit-gold respectively before examining in TEM. A)LF10 and B)slpA9

4 Discussion

4.1 Surface proteins in Gram positive bacteria

The surface of each microbial pathogen forms the frontier for encounter with the host and it is therefore inevitable that bacterial proteins in this location possess a variety of functions of significance to interaction with the host. The nature of these surface proteins and their activities can be exploited for identification and characterisation of the pathogen, the detection of the pathogen's presence in the host, the development of antimicrobials and vaccines. Given the importance of their location, the functions of bacterial surface proteins are diverse, ranging between factors required for growth and cell division, to adherence, colonisation and the formation of biofilms, to lytic enzymes required for dissemination in the host, and toxins.

There are different types of systems in Gram positive bacteria to deliver these proteins to the surface. Proteins can be targeted to the cytoplasmic membrane through possession of an amino terminal signal peptide, a defining feature of one transport pathway. Most surface proteins are transferred in an unfolded state from the cytoplasmic side of the membrane to the outer side *via* this process, known as the General Secretory Pathway, or Sec-dependent pathway because of the contribution of proteins (SecY, E, G, F, F and A) that assemble into channel and ATP-driven motor complexes. In contrast, the TAT (twine argentine) pathway is involved in the transport of proteins across the membrane in a folded state. Other protein transportation systems have also been described such as ATP- binding transporters employed in the export of bacteriocins or holin- like proteins (Scott and Barnett, 2006).

If proteins that are transported *via* these systems are destined for attachment at the bacterial surface, covalent and non-covalent attachment may take place. In covalent attachment, a transpeptidase termed a "sortase" may attach the protein to peptidoglycan, exploiting conserved carboxy-terminal features. The enzyme was

originally identified in *Staphylococcus aureus* (Mazmanian et al., 1999) but enzymes with similar activities have been discovered in a variety of Gram positive bacteria such as *Listeria monocytogenes* (Bierne et al., 2002), *Bacillus anthracis* (Gaspar et al., 2005) and streptococcal species (Barnett and Scott, 2002, Bolken et al., 2001). Generally the protein substrate carries a carboxy-terminal anchoring domain consisting of an LPXTG motif that can be recognised by sortase which then carries out transpeptidation at the threonine residue. Other forms of surface attachment can take place by transglycosylation and transpeptidation. Another feature of the action of sortase A is in the polymerisation of pilli as observed in *Streptococcus agalactiae* (Lauer et al., 2005). Reports have shown that sortases other than sortase A exist with different substrate sequences such as NPQTN for sortase B (Mazmanian et al., 2002) in *S. aureus* and QVPTGV for sortase C2 from *S. pyogenes* (Barnett et al., 2004) .

Non-covalent interactions can also bind proteins to the surface of Gram positive bacteria. Binding to teichoic or lipoteichoic acid is one such process in which repeats at the carboxy-terminus of the surface protein can mediate attachment to the cell wall. PspA and LytA are two examples of choline binding proteins from *S. pneumoniae* that attach in this way (Briles et al., 1998, Holtje and Tomasz, 1975). In other examples, glycine-tryptophan rich (GW repeats) as found in InlB from *L. monocytogenes* mediate binding to lipoteichoic acid (Gaillard et al., 1991). The other group of highly abundant proteins that attach non-covalently to the surface are the surface layer proteins. In these examples, the amino-terminus of the proteins form a domain that is involved in non-covalent attachment (Etienne-Toumelin et al., 1995). The LysM motif is a further example of a discrete protein domain which is responsible for non-covalent binding of proteins on the surface of Gram positive bacteria. The major autolysin (AcmA) of *Lactococcus lactis* is one example of a protein that contains this domain, essential for attachment to the surface (Steen et al., 2003).

4.2 Surface proteins of *C. difficile*

The large glucosyltransferase toxins of *C. difficile* and the ribosyltransferase binary toxin produced by some strains are known to play important roles in the pathogenesis of infection. However, the surface proteins of vegetative cells have also attracted attention in attempts to clarify their roles and their functions in the pathogenic process.

Proteomic studies have used different methods to extract proteins from the bacterial surface and have documented a wide range of around 49 components. Although many have proved to be paralogs of SlpA, the main component of the surface layer, others have been related to the flagellum, proteins with enzymic activities, transporters, and proteins elicited by heat or cold (Wright et al., 2005). These surface proteins are likely to contribute many different functions to the biology of *C. difficile* but amongst them, a group of 29 have been classified as a family through the possession of a putative cell wall binding domain. In addition to members of this family, other proteins are also considered very likely to be components of the bacterial surface. A sortase has been identified along with seven putative sortase substrates; one of these is CD0384, a collagen binding protein, and another is CD3246 (<http://nihserver.mbi.ucla.edu/Sortase>); (Fagan et al.). Other proteins have been known or suspected components of the surface of *C. difficile* for some time. Examples include Cwp66 (Waligora et al., 2001), Cwp2 and CwpV (Emerson et al., 2009), Cwp84 (Chapeton Montes et al.), Fbp68 (Hennequin et al., 2003), GroEL (Hennequin et al., 2001b) and the proteins that comprise the flagellum (Pechine et al., 2005b, Pechine et al., 2005a).

The extraction of these proteins from the bacterial surface so that their properties and contribution to the biology of *C. difficile* might be studied is difficult if not impossible. Therefore a number of known and putative surface proteins were chosen for study according using data from the *C. difficile* 630 genome project, the coding sequences were cloned and the target proteins were expressed in *E. coli* using a LIC vector system from Novagen. The underlying aim of this part of

the project was to gain access to purified recombinant proteins which could be used for the isolation of antibodies by phage display. The target sequences excluding signal sequences were amplified and ligated into the pET-32 EK/LIC vector. The cloned sequences were confirmed by sequencing into the inserts from either terminus using primers that annealed to vector-encoded sequences. Fusion of a His-tag to each recombinant target enabled the expressed proteins to be purified from cell lysates by affinity chromatography. Whilst the molecular weight of the purified proteins and their reaction with antibodies against a fused detection tag were indicative of their likely identity, mass spectrometry was used for confirmation. Other groups have also followed this approach in attempting to characterise the properties of proteins from the surface of *C. difficile* and other virulence factors.

Codon usage in *C. difficile* can present difficulties when attempting to express native clostridial sequences in bacterial systems such as *E. coli*. That said, a recent study reports the expression and purification of recombinant Cwp19 for structural investigation and showed that the *E. coli* expression vector pET28a could be used successfully. The recombinant Cwp19 carried a His-tag for purification by affinity chromatography. The study showed successful purification and crystallisation of Cwp19, a putative glycosyl hydrolase with amino-terminal domains likely to be responsible for attachment to the clostridial cell wall (Kirby et al.)

GroEL from *C. difficile* has also been successfully expressed in *E. coli* by cloning into vector pGEX-6p-1 to prepare a recombinant fusion to glutathione S-transferase. This recombinant protein was purified through a glutathione-Sepharose-4B column for the preparation of antibodies. These were used to identify the location of GroEL in the clostridial cell and to document its release under heat shock (Hennequin et al., 2001b) project objectives similar to those posed in this study. The peptidoglycan hydrolase (Acd) was also cloned and expressed in *E. coli* using a further expression system, the pQE-32 vector. Although the expressed protein was purified by a nickel chelation chromatography under denaturing

condition (8M urea) (Dhalluin et al., 2005) later functional studies showed that its ability to hydrolyse the bonds between N-acetylglucosamine and N-acetylmuramic acid in peptidoglycan was unimpaired, thereby confirming Acd's activity as an N-acetylglucosaminidase. Amongst those cell surface proteins thought to make a direct contribution to the virulence of *C. difficile*, a 68 kDa fibronectin binding protein, highly conserved between clostridial isolates, has been expressed in *E. coli* as a glutathione S-transferase fusion protein using again the pGEX-6p-1 vector. Mirroring the work of their earlier studies with GroEL, Hennequin and colleagues expressed the fusion protein in *E. coli*, purified it through a glutathione Sepharose 4B column and raised specific anti-Fbp68 antibodies.

Using these reagents, the authors were able to show that this protein is localised on the surface and plays a role as a bridge between bacteria and the host cell. A competition inhibition assay with anti-Fbp68 antibodies revealed that the target protein can bind soluble and immobilised fibronectin (Hennequin et al., 2003). Again, the study described in this thesis shares a common strategy – purification of recombinant target protein for the isolation of specific antibodies – with this published work. Based on a further study, fibronectin binding protein is regarded as one of the adhesion factors of this pathogen, a factor that is stabilised by the presence of manganese (Lin et al.). A recently released report tackles this issue directly by construction of a mutant through Clostron technology. *In vitro* tests revealed that mutant bacteria deficient in the production of Fbp68 could adhere to cultured epithelial cells at a higher rate than wild type *C. difficile*. In contrast, studies of caecal colonisation and intestinal implantation showed a slower rate of colonisation for the mutant *versus* wild type (Barketi-Klai et al.). Proteins may have different roles in adherence and colonisation at the same time which should be considered. One obvious feature of the surface composition of *C. difficile* is the presence of the flagellum. In many bacterial pathogens, motility plays an essential role in virulence and in some cases, flagella are thought to serve as adhesins. As an example, deletion of *flgE* of *Vibrio vulnificus* revealed decreasing mobility and adhesion of this pathogen to host cells (Lee et al., 2004). Another case is

Pseudomonas aeruginosa where motility by flagella is a preliminary step for infection of the respiratory tract and stimulation of immune system through mucosal epithelial cells. Due to this pivotal role in pathogenesis, flagella of this bacteria are considered as an important factor in vaccine development (Prince, 2006).

In *C. difficile*, the contribution of flagella has been investigated and it is known that the flagellae are immunogenic in human patients. In one study, flagella were isolated from overnight culture on a solid medium and showed a positive signal in immunoblots using rabbit antiserum (Delmee et al., 1990). Other investigators have used molecular methods to prepare parts of the flagellum by recombinant methods. FliD was cloned into an expression vector (pGEX-6p-1), producing a GST-FliD fusion protein on induction. The protein appeared able to mediate adherence to mucus or cell receptors (Tasteyre et al., 2001b). In another study, FliC was also prepared *via* the same expression vector. The resulting protein possessed a molecular weight of 39 kDa protein but the size of the FliC proteins differed between strains according to the length and sequence of the *fliC* gene (Tasteyre et al., 2000a). In the same study, antibodies raised against FliC were used to localise flagella on the bacterial surface using immunogold methods, a parallel with the aims of this project.

Other unexpected findings have emerged from work on Cwp84, a cysteine protease. The gene lies just downstream of *cwp66* gene and both are paralogs of *sfpA*, carrying cell wall-binding PF04122 motif (Sebahia et al., 2006). Based on its gene sequence, the mass of the Cwp84 protein is 87.3 kDa and it is thought to carry a peptidase domain at its amino-terminus and three cell wall binding domains at the other end (Fagan et al.). Perhaps because of its proteolytic activity, expression and purification of a recombinant GST-Cwp84 protein has not been successful, a problem shared with other cysteine proteases (Wandersman, 1989). Like cysteine proteases such as Spe from *S. pyogenes* (Doran et al., 1999, Musser et al., 1991), Cwp84 is very conserved between different strains and its activity can

be detected in the early exponential phase. This finding strengthened the idea that Cwp84 might play an important role in the maturation of other cell wall proteins (Savariau-Lacomme et al., 2003) as already shown for SpeB in *S. pyogenes* (Berge and Bjorck, 1995). The inability to purify GST-Cwp84 was overcome by expressing *cwp84* in pET-28a and purifying the recombinant enzyme using the attached His-tag. Automaturation, a mechanism that is seen for several cysteine proteases, was confirmed under reducing conditions resulting in a protein of 61 kDa.

This automaturation was previously reported for other bacterial proteases (Doran et al., 1999). Proteolytic activity against host proteins located in the extracellular environment – molecules such as fibronectin, laminin and vitronectin – was demonstrated *in vitro* and could be neutralised with antibodies against Cwp84 or chemical inhibitors and investigators proposed that the action of the protease on host tissue might facilitate the actions of toxins produced by *C. difficile* (Janoir et al., 2007). Given the important contributions proposed for Cwp84 in cell wall assembly and pathogenesis, gene knockout results were unexpected: mutation of *cwp84* prevented the maturation of SlpA creating a change in colony morphology that was evident when compared with wild type. Mutants grew more slowly than wild type in liquid medium. However, experiments in animal models showed that there was no difference in virulence between mutants and wild type (Kirby et al., 2009). The putative protease CD1751, another enzyme likely to be associated with the bacterial surface and possessing similarity with Cwp84, had no effect on processing of SlpA and genome analysis of the virulent strain R20291 revealed a deletion of the gene.

To understand the localisation of Cwp84, another study prepared different mutant forms of this protein by expression in pET28a. Investigation revealed that the fully processed form of Cwp84, a protein cleaved at Lys 91 and lacking proteolytic activity, could associate with the bacterial cell wall (Chapeton Montes et al.). Other recent data has suggested contribution from another cysteine protease (Cwp13) in the sequential processing of Cwp84 processing (de la Riva et

al.). Immunisation of experimental animals with Cwp84 appears to reduce intestinal colonisation and the severity of infection but fails to provide complete protection (Pechine et al., 2011). Understanding of the role of Cwp84 in the biology of *C. difficile* and its virulence is incomplete.

The most abundant protein on the surface of *C. difficile* is SlpA. This undergoes post translational processing from a single gene product to create a low molecular weight (LMW) amino-terminal protein and a high molecular weight (HMW) product derived from the remainder of the initial translation product (Calabi et al., 2001). This surface protein has been isolated by chemical extraction methods or by expression as a recombinant product, allowing investigators to study characteristics such as sequence variability between strains (Calabi and Fairweather, 2002), the role of SlpA in binding *C. difficile* to the gastrointestinal tissue of the host (Calabi et al., 2002) or the interaction between LMW and HMW components in formation of the cell envelope (Fagan et al., 2009).

One of the first reports on the properties of SlpA used EDTA extraction combined with gel filtration and ion exchange chromatography to purify the LMW protein. A 36 kDa protein was obtained that lacked protease activity but reacted with the sera from patients infected with *C. difficile* (Cerquetti et al., 1992b). In another approach to extraction, 8M urea generated a crude preparation (Cerquetti et al., 2000a) that could be purified by anion exchange chromatography (O'Brien et al., 2005). Other investigators have used prolonged exposure to 5M guanidine hydrochloride to extract SlpA from bacterial pellets. This was dialysed before further investigation (McCoubrey and Poxton, 2001, Sanchez-Hurtado and Poxton, 2008). Extraction with 0.2M glycine has also been explored, extracts being neutralised with 2M Tris before further analysis (Calabi et al., 2001, Qazi et al., 2009, Drudy et al., 2004). In a comparison of a number of different approaches with two-dimensional gel analysis, Wright *et al.* found glycine extraction to be more efficient in recovering SlpA (Wright et al., 2005) and given the adoption of this approach by the Imperial College group in studies of SlpA and other surface

proteins from *C. difficile*, it was used in the present study. There are some reports of the use of molecular methods to prepare SlpA as a recombinant protein and sequencing of the gene from different strains in typing scheme or studies of strain variation (Eidhin et al., 2006, Karjalainen et al., 2002). pET28 has been used as an expression vector for the production of LMW and HMW components of SlpA allowing purification of the recombinant products (Calabi et al., 2001, Calabi et al., 2002). The LMW was prepared as a GST-LMW fusion protein *via* insertion of amplified sequence into pGEX-6P-1 and purification through glutathione-Sepahrose 4B (Brun et al., 2008). Given these details, surface proteins were considered as the targets for selection of recombinant antibody from an antibody phage display library.

4.3 Recombinant antibodies and their interaction with different surface proteins

Phage display libraries have been used increasingly in a vast range of biological fields. Antibodies selected from these libraries can be applicable in therapy or diagnostics (Van Bockstaele et al., 2009, Wesolowski et al., 2009) or to better understand the nature of basic biological processes or disease.

In virology there are many reports of the use of scFv antibodies across a variety of fields particularly in the isolation and development of antibodies with antiviral activity (Parren and Burton, 2001). In one example, antibodies against HIV were selected from a phage display library that was constructed from patient with lupus. These experiments isolated binders against residues 421-436 of gp120 showing that patient samples can be a valuable source of neutralising antibodies against HIV (Karle et al., 2004). The approach can use used for a range of hosts – for example, antibodies against foot and mouth disease virus were isolated from the bovine repertoire (Kim et al., 2004) – and adapted for use with conventional rodent monoclonal antibodies (ShengFeng et al., 2003).

In bacteriology, phage display technology has also been used to enormous effect in selecting antibodies from synthetic libraries or libraries constructed from immunised or infected animals. In many cases, antibodies can be isolated that possess biological activity against the target molecule. The alpha toxin from *Clostridium perfringens* was used as target for a library that was made from a hybridoma cell line producing monoclonal antibody. Experiments successfully isolated a scFv that could be expressed as a recombinant protein with neutralising activity against the phospholipase activity of alpha toxin (Zhao and Xu, 2001).

Pertussis toxin was another target used to generate specific recombinant antibodies with neutralising activity. The selected antibodies against this toxin had the capability to recognise and also neutralise pertussis toxin in assays using CHO cells. The source of the scFvs in this study was a library produced from the peripheral blood of a patient infected with *Bordetella pertussis* (Williamson and Matthews, 1999). It has also been reported that phage display using an immune library yielded scFv against a range of targets from *Mycobacterium tuberculosis* but managed to avoid isolating binders to the immunodominant 65 kDa antigen. This finding demonstrated the potential of phage display technologies to generate diverse populations of reactive binders that are of value in documenting the immune response to infection and vaccination (Cummings et al., 1998).

Further advantages are shown through work with a hybridoma clone able to produce high affinity monoclonal antibody for the detection of Staphylococcal Enterotoxin B. In this example, the hybridoma cells had lost the ability to secrete antibody as a consequence of repeated culture. A phage display library was constructed by transferring the immunoglobulin genes into a phagemid vector. Results showed that the anti-toxin reagent could be rescued and could be used successfully in assays with a high rate of success (Singh et al.).

With regard to Gram positive pathogens, investigators have reported the isolation of two scFvs recognising the protective antigen of anthrax using a library

made from immunised chimpanzees. Antibodies showed the ability to neutralise the cytotoxicity of lethal toxin from *Bacillus anthracis*. Further analysis showed that these two antibodies bound to two conformational epitopes at the carboxy-terminus of protective antigen (Chen et al., 2006). In another study, a recombinant antibody against *Bacillus anthracis* spores was characterised. The antibody showed no cross reactivity with spores from related species of *Bacillus* and with high affinity, they can be used for the detection of anthracis spores in immunoassays (Mechaly et al., 2008).

In other studies, immunised and non-immunised libraries were used for isolation of antibodies against botulinum neurotoxin type A. Epitope mapping for scFvs selected from the immunised library showed differences between antibodies derived from each library and significantly, neutralising antibodies that were isolated from the immunised library were not found amongst the many anti-toxin scFvs from the non-immune library. In addition, this report revealed that mixing of individual scFvs recognising different epitopes, improved the potency in neutralisation assays when compared to the biological activity of the individual binders (Amersdorfer et al., 2002).

A recent study reports the use of tandemly linked scFvs (chelating recombinant antibodies) as an effective approach in increasing the affinity of binders to non-overlapping epitopes on protein targets. In other studies using virulence factors from Gram positive bacteria, scFvs were isolated against different parts of tetanus toxin and it was shown that some were able to neutralise toxin activity by interference with ganglioside binding. Interestingly, scFvs with neutralisation activity could be identified that did not bind this suggests that scFvs able to inhibit the biological activity of bacterial virulence factors may be able to work through interference with conformational change in the target (Scott et al.).

From all of the above examples, it is apparent that phage display offers researchers convenient access to powerful reagents for analysis of bacterial

virulence. This is the first report of the selection of recombinant antibodies against a range of surface proteins from *C. difficile*. As explained already, several targets in this location are known to contribute to the pathogenic process. The project therefore aimed to express a range of targets – some of known location and function, others more uncertain – in *E. coli* and purify the recombinant proteins for selection of scFv antibodies from the semi-synthetic Tomlinson phage display libraries. Although there are reports of the isolation of antibodies against some of these targets from *C. difficile* so that the location and function of the clostridial proteins can be understood better, this the first report of the application of antibody phage display in this context.

The Tomlinson I and J libraries differ in the manner of their diversification. Diversity in library I was carried by incorporation of DVT codons (D: A, G or T; V: A, C or G) at selected positions in the scFv CDRs. The DVT approach provides nine possible nucleotide combinations without the possibility of including a stop codon. Hence, the capacity for diversification is modest but a strength is that stop codons are excluded. In contrast, library J was diversified by incorporation of the NNK triplet (N: any nucleotide; K: G or T) at the same locations in the scFv reading frame.

Whilst the capacity for diversity is greater – 32 possible combinations – one stop codon can be incorporated. Propagation of the library can deplete its diversity – some clones may replicate or infect the bacterial host with greater success than others – and based on experience, it would be wise to check the diversity of libraries like this before starting selection. Initial aliquots from the library failed to show evidence of successful selection resulting in wasted time. Other aliquots of the library were checked in trial selections against BSA and once evidence was seen of positive selection, experiments began with the clostridial target proteins.

These reagents were used for selection on all purified target proteins. Over three rounds of selection, polyclonal phage ELISAs showed sharp increases in

absorbance after the second round and flattened at the third round. Phage recoveries were calculated and the results revealed enhanced yields towards the last round of selection, the extent of this rise varying between different targets. This could be interpreted to mean higher representation of specific binders round on round, elevation in the affinity of dominant clones for the target or as recoveries stabilise, equilibration of selective effects between specific and non-specific interactions. Other studies from this laboratory have observed similar effects (Li et al., 2003).

While some phage display experiments increase the number of rounds of selection from 3 to 5, this does not necessarily increase the specificity or quality of the recovered antibodies. In this study, phage taken from the third round of selection were checked in ELISA for reaction with BSA and skimmed milk but reaction with these blocking reagents were not detected and a constant base line absorbance was measured with increasing numbers of phage in the assay. The clones chosen by monoclonal phage ELISA were expressed as soluble scFvs and retested. Positive results were obtained in all cases though the strength of ELISA signal was variable. Those soluble antibodies that gave strongest signals in ELISA were chosen for further study but again, they showed no reaction with BSA or skimmed milk.

Some reports note differences in frequency of positive reaction in ELISA with phage *versus* soluble antibodies. This might arise from at least two causes. The first are avidity effects. M13 phage is a filamentous phage that can possess up to three copies of the scFv-pIII fusion protein at its surface. The presence of three copies of the fusion protein can create multivalent binding to target, a feature that gives rise to high avidity phage antibodies (de Wildt et al., 2002). This has been reported by investigators who have used the Tomlinson antibody libraries (Wu et al., 2007). Under these circumstances, the phage are able to bind to target more strongly than when monovalent scFv is expressed or drive non-specific interactions with the selecting surface rather than the target itself. There are different ways to

solve these problems such as changing the composition of the immunotubes between rounds of selection, alternating blocking solutions in cycles of panning, or increasing the detergent concentration in the washing steps after each round.

The second cause of differences between phage and soluble scFv ELISAs for a given clone can be the presence of TAG stop codons that are suppressed in *E. coli* TG1 during the production of phage, but which result in premature termination in *E. coli* HB2151 when soluble scFvs are expressed. The NNK diversification strategy used in construction of the Tomlinson J library will incorporate TAG with approximately 3% frequency at each diversified codon. The stability of scFv could also be a factor (Lorimer et al., 1996).

None of these issues proved to be complicating factors in the present study and in summary, phage display provided a fast, efficient and productive strategy for isolation of monoclonal antibodies against a range of proteins from *C. difficile*. It also allowed the screening and analysis of significant numbers of clones. From an initial panel of 672 soluble antibodies against a range of targets, 100 were picked on the basis of the strength of ELISA signal and from these, 32 were chosen to comprise the final panel of reagents. This process of screening used soluble scFvs against recombinant targets in dot blots and Western blots. The reactivity of the scFvs in these analyses suggests that the epitopes recognised on each target protein may be linear rather than conformational. DNA sequencing showed that for a given target, CDR sequences were frequently (but not always) unique. Of the scFvs isolated, only those against the putative sortase target carried identical sequence in all CDRs with repeating hydrophobic and hydrophilic amino acids, indicating reaction with a single epitope on the sortase and perhaps the emergence of a single, dominant clones through the selection process. Identity was noted for 2 of the 3 scFvs against the full-length Cwp66 and interestingly, one of the scFvs against the amino-terminal region of Cwp66 possessed the same CDR sequences. No such similarity was noted for scFvs against full-length Cwp66 and the carboxy-terminal region of the protein. Hence, specific selection was easily

achieved but it was also clear that single clones failed to dominate the selection process.

The properties of selected scFvs were assessed through a number of experimental approaches using the targets from bacteria. One approach was to extract the target from bacterial cells and to assess if the scFvs could recognise the protein in Western blots. Analysis of selected scFvs against SlpA in this way revealed that the antibodies could bind to extracts of the *C. difficile* strains 630 and M120 when extracts were treated under denaturing and non-denaturing conditions but that the scFvs failed to bind to extracts from strain R20291. Firstly, this suggests that the epitopes for this panel of antibodies are probably linear and hence treatment with SDS does not disrupt the sites of interaction. There are no predictive methods for identification of linear or conformational epitopes and it should be recalled that targets in the context of this project are not “antigens” in the usual immunological sense: the libraries simply contain a wide range of CDR sequences and their composition has not been influenced by exposure to foreign materials unlike the context *in vivo*. For example, a panel of antibodies against severe acute respiratory syndrome (SARS) revealed both linear and conformational epitopes in the viral S protein (Chou et al., 2005) and the similar effects can be seen in antigenic site A of the haemagglutinin of influenza virus (Ohkura et al.).

Secondly, the results suggest that the epitopes recognised by these antibodies may be shared between the SlpA molecules of 630 and M120; as explained below, this observation needs careful consideration. That aside, it is clear that the epitopes recognised by the anti-SlpA scFvs are not shared between 630 and R20291. Protein alignment between the SlpA sequences of 630 and R20291 has revealed that the cleavage site between the LMW and HMW components of SlpA are the same but to either side, there are differences in sequence.

The extent of this variation is higher in the LMW than HMW. Conservation of the sequences around the SlpA cleavage site has been reported for other strains of *C. difficile* (Calabi et al., 2001) and 630 is substantially similar to R20291 at sites thought to mediate interaction between LMW and HMW. The likelihood is therefore high that the binding sites for anti-SlpA scFvs tested in these experiments lie elsewhere. Potentially, the location of epitopes could be refined by surveying a broad range of strains of *C. difficile* and through exchange of sequences between reactive and non-reactive SlpA proteins. The ability to express *slpA* in *E. coli* makes these experiments easy to conduct and the availability of anti-SlpA antibodies that are strain specific might assist in monitoring the spread of strains of *C. difficile* and the emergence of new strains in the hospital environment.

The scFvs developed in the project may also contribute to a better understanding of the structure and function of SlpA. A conservation of sequence has been noted in the amino and carboxy-termini of LMW amongst different strains suggesting that together, these sequences form a distinct region – domain 1 – that is internal to the surface layer complex. A separate region of LMW is thought to be external – domain 2 – and shows extensive strain to strain variability (Fagan et al., 2009). If this model is correct, it would seem unlikely that scFvs against sequences against domain 1 would have been isolated since the panning process used SlpA that was in native form. Others have reported that the HMW component of SlpA from *C. difficile* may have a role in binding the bacteria to host proteins like collagen I and that bacterial adhesion can be inhibited with a specific antiserum (Calabi et al., 2002). Given the ease with which anti-SlpA scFvs were isolated in this project, one future direction would be to seek anti-HMW antibodies and test if they possess similar properties.

Antibody levels, including levels of IgM, against this surface protein are thought to be an important factor in preventing relapse and strain specific responses are also considered important (Drudy et al., 2004). The IgG responses of patients against SlpA (Pantosti et al., 1989b) and the contribution of these responses to the

adaptive immune response have suggested that multicomponent vaccines against *C. difficile* might have value for patients at risk (Ausiello et al., 2006), such vaccines perhaps including a mixture of toxin A components and the low molecular part of SlpA since this appears to enhance the immunogenicity of co-administered antigens (Brun et al., 2008). These studies and others illustrate the importance of a more detailed understanding of the structure, function and biological activity of SlpA.

Results obtained with extracts from *C. difficile* M120 need careful consideration: nearly all of the scFvs used to probe Western blots appeared to be able to bind to a protein of a size consistent with the LMW of SlpA from the M120 strain. This applied to scFvs directed against GroEL, Sortase B, FliC and CspA. Notably, anti-CspA scFvs recognised the recombinant clostridial protein on Western blots but failed to bind to proteins in extracts from *C. difficile* strain 630 or R20291. This might be for a variety of reasons – for example, CspA might not be present in bacterial lysates at a concentration sufficient to allow detection – but reaction with a 36 kDa protein from M120 was striking. Given the number of scFvs that appear to react with this protein, tentatively, LMW from SlpA of M120, the most likely explanation is that it is able to bind scFvs not through interaction with residues in the CDRS of the antibodies, but through other parts of the scFvs.

There appear to be no reports of proteins from *C. difficile* with the general ability to bind immunoglobulin although this aspect of the pathogenic process is seen in many other pathogens. Isolated reports suggest that, in contrast, toxin A may be able to bind host antibodies through its ability to interact with carbohydrates (Cooke and Borriello, 1998) and this may extend to other proteins associated with immunoglobulins (Dallas and Rolfe, 1998). From the host perspective, interaction between GALT and the intestinal microflora has profound effects on development of the immune system and some VH sequences are known to possess ligand binding sites for endogenous B cell superantigens (Rhee et al., 2005) or superantigens such as Protein A that are derived from bacteria (Rhee et

al., 2004). Recent studies with *Bacillus subtilis* and other species (Severson et al., 2010) have shown that spore proteins including ExsK are able to interact with immunoglobulins from some animals, and that scFvs generated from certain VH variants retain the ability to interact with these immunoglobulin binding proteins (Severson et al.).

The concept that some LMW proteins from *C. difficile* may possess the ability to bind host antibodies through superantigen-like activity may alter current views of the immunogenic properties of the LMW (Spigaglia et al.). Currently, it is recognised that many components of the clostridial cell wall are immunogenic and that sera from patients are reactive with native or recombinant forms of these proteins. For example, the sera from different patients have been tested in Western blot following two dimensional electrophoresis, with additional mass analysis to identify those spots that were reactive (Wright et al., 2008). In all, 42 proteins were reactive in this way, of which 11 were cell wall or membrane associated. However, across the panel of patient sera only 3 reacted with Cwp84, and 4 reacted with FliC. The immunodominance of SlpA emerged from the finding that all sera were reactive with components of SlpA, mostly LMW (Wright et al., 2008).

As described earlier, antibodies against bacterial virulence factors sometime possess the capacity to block the biological activity of these proteins. For some of the clostridial targets used in this study, function remains uncertain (*eg* CspA is thought to be a protein produced in response to cold shock but its role in the biology of *C. difficile* is unknown). For others it is difficult or impossible to assay the biological activity of the clostridial target (*eg in vitro* assays of sortase activity exist (Aulabaugh et al., 2007) but they require specialist equipment and expertise beyond the range possible in this project).

However, it was possible to assess whether the panel of scFvs against FliC and FliD were able to inhibit bacterial motility, using a scFv against LMW, a target

known to be at the surface of *C. difficile* as a control. In contrast to other scFvs tested in Western blotting, the scFvs against FliD did not generate a signal in blots when they were used as probes for flagellar extracts from strains 630 and R20291. This might have arisen because of the low quantities of FliD in the extracts. It is unlikely that the epitopes recognised by these scFvs are conformational as the antibodies recognised recombinant FliD after SDS PAGE and blotting. Antibodies against FliC reacted strongly in blots prepared with flagellar extracts from 630 but weak reaction was evident with material from R20291.

Protein alignment for FliC and FliD sequences from the two strains showed a high degree of identity for both proteins with greater variation in the central region of each protein. This alignment and the results from blotting suggest that the binding site for the anti-FliC scFvs may lie in the diverse central region. Others have reported sequence variation in this part of FliC among a wider range of different strains of *C. difficile* (Tasteyre et al., 2000b). The authors reported that FliC from six strains reacted with a polyclonal anti-FliC antibody (Tasteyre et al., 2000a), noting that the predicted molecular weight from sequencing of the gene was different to the molecular weight of FliC from bacterial extracts. More recent bioinformatics analysis of different strains has shown that the protein undergoes glycosylation, the differences in glycan biosynthesis genes further impacting upon the masses observed for FliC from different strains (Twine et al., 2009). In the cited experiments, bacterial motility was assessed by visualisation of growth and spread in soft agar. Aside from motility, FliC and FliD are thought to contribute to bacterial adherence to intestinal mucous *in vitro* and to the mouse intestine *in vivo* (Tasteyre et al., 2001a). Interestingly, bacterial mutants with complete loss of these proteins were more virulence and were enhanced over wild type in their ability to colonise the hamster intestine (Dingle et al.).

In the present study, the biological activity of anti-flagellar scFvs was assessed by binding antibodies to *C. difficile* 630 and examining bacterial motility in a simple, soft agar assay. Experiments showed that some but not all anti-FliC and

anti-FliD scFvs possessed the capacity to inhibit bacterial motility. Sequence data was only available for two scFvs against FliC and both were able to neutralise flagellar activity. Of the three anti-FliD scFvs that were sequenced, none carried identical CDR sequences in the heavy or light chain domains but there were some similarities between clones N6A4 and N6B5. While the former antibody only achieved partial inhibition of bacterial motility, the latter was more effective. It is therefore unclear whether the binding of scFvs to particular regions of the flagellar proteins is the key determinant in their biological activity, whether the affinity of the interaction explains partial from complete effect, or whether other factors are at work. The mode of action is similarly uncertain.

Some studies describe the immunisation with flagella or their constituent proteins and effects on motility (for example, Faezi et al., 2011) but the use of native, multivalent antibodies might lead to aggregation of flagella at the bacterial surface or agglutination. There are very few published studies on the impact of recombinant antibodies on flagellar activity. In one report, scFvs were generated against the H7 flagella of *E. coli* O157 (Kanitpun et al., 2004). Properties included the ability to act in competition with polyclonal antibodies and thereby to inhibit bacterial agglutination. The authors propose the scFv might be of value in a capture assay for *E. coli* O157. In another study, scFvs were generated from conventional monoclonal antibodies against an immunodominant surface protein from the plant pathogen *Spiroplasma citri* (Malembic et al., 2002). The scFvs were able to inhibit bacterial growth and motility but as their target, spiralin, appears non-essential for motility (Duret et al., 2003), there are significant differences with the current study.

Whilst the role of flagella in pathogenesis has yet to be resolved, it is clear that these features of the cell are immunogenic as antibodies from patients infected with *C. difficile* are frequently reactive with FliC and FliD (Pechine et al., 2005a). When rates vary for FliC, this is most likely due to the variability of sequence of those parts of the protein that are surface exposed. Other studies have

documented the recognition of clostridial surface proteins by sera from patients. The surface exposure of other proteins such as Cwp66 (Waligora et al., 2001) also gives rise to high rates of seropositive reaction and many patients develop antibody against surface exposed domain of the protein (13 out of 17 samples) in contrast to recognition of the amino-terminal regions (2 out of 17 sera). Cwp84 is also recognised commonly during infection (Pechine et al., 2005b).

These observations have prompted investigators to consider whether proteins at the bacterial surface might contribute to a vaccine against *C. difficile* infection. In experiments with mice, immunization with FliD *via* different routes showed the best IgG responses were achieved by combined intravenous and subcutaneous delivery whereas the rectal route was the best mucosal immunization route. This provoked better IgA responses than intranasal delivery. These studies showed that the lowest rates of colonisation with *C. difficile* were achieved by rectal immunization with cell wall extracts and a cholera toxin adjuvant, but that a mixture of FliD and Cwp84 came close in effectiveness (Pechine et al., 2007).

Subsequent studies have confirmed the significance of host responses against Cwp84 but noted that immunization with this protein only has partial effect on colonisation and protection (Pechine et al., 2011). Vaccination of experimental animals with SlpA has similarly incomplete protective activity (Eidhin et al., 2008). In contrast, the stimulation of neutralising antibodies against the protein toxins of *C. difficile* is seen as a very promising way to prevent or treat infection and neutralising IgG and sIgA are thought to have a critical role in regulating the severity of infection (Hussack and Tanha). Different vaccines have been under development for the stimulation of anti-toxin antibodies (Giannasca and Warny, 2004a) but the use of natural or recombinant antibodies as passive immunotherapeutics is also under development, building on early observations that protection could be transferred through milk from vaccinated adult hamsters to their offspring (Kim and Rolfe, 1989). One example from the literature describes

the selection of recombinant antibodies against toxin B of *C.difficile* (Deng et al., 2003) using a hyperimmunised scFv library as source for selection.

In many different bacteria, GroEL is recognised as a chaperone involved in ensuring the proper folding of proteins (Zeilstra-Ryalls et al., 1991). The over-expression of GroEL has also assisted in generating high yields of recombinant proteins in heterologous expression systems (Arbabi-Ghahroudi et al., 2005) and in the production of recombinant antibodies, co-expression of antibody and chaperones has been found to improve the functional production of antibodies by 2.4-fold (Maeng et al.). With this in mind, the up-regulation of GroEL and other clostridial proteins in response to heat might be expected; its appearance at the surface and its proposed role in adherence may seem surprising (Hennequin et al., 2001) but a body of evidence indicates GroEL serves these functions in other species (reviewed in Hennequin et al., 2001).

Consistent with its function, alignment of GroEL protein sequences from *C. difficile* 630 and R20291 and BLAST analysis showed high degrees of similarity between strains and the presence of conserved domains. In this study, the scFvs isolated against GroEL appeared to recognise the native protein, a protein of 60 kDa, in Western blots using extracts from three bacterial strains (630, M120 and R20291), after heat-shock of cultures. Although scFv N5E10 reacted with a protein of 60 kDa from all three bacterial strains reaction with a protein likely to be the LMW of M120 was also evident. This observation was also made with other scFvs against the putative sortase and CspA as discussed earlier. While heat shock seemed to elicit production of GroEL in *C. difficile*, allowing detection with the scFvs, antibodies against Cwp66 and CspA were unable to detect the presence of their native targets in bacterial extracts. It may be that the quantities present were insufficient for detection or that as the expression these proteins required specified conditions like temperature that were not met by culture conditions. Protein alignment and BLAST analysis showed quite similar sequences for the

putative sortase in strains 630 and R20291 and high similarity with the sortase from *Turicibacter sanguinis* (Cuiv et al.).

At the outset of the project, it was hoped that scFvs against the known and putative surface proteins of *C. difficile* could be used in immunofluorescence and electron microscopy to establish the location and distribution of the targets, and allow a simple screen for conditions that might regulate their expression. Although purified scFvs against FliC, FliD, LMW and HMW seemed able to bind to their targets in immunofluorescence experiments, a degree of caution needs to be exercised given that for strain M120, binding of scFvs through target-independent processes was evident in blotting. Other studies with antisera against the LMW of SlpA convincing contrast between immune and non-immune sera (Cerquetti et al., 1992a) and immunofluorescence microscopy has also assisted the characterisation of the fibronectin binding protein from *C. difficile* (Hennequin et al., 2003).

As a system, immunostaining with scFvs has the potential to assist in validation of drug targets and the development of new therapeutics (Laforce-Nesbitt et al., 2008) Pongpair et al., 2009), aspects that are important for infection with *C. difficile* but also more widely, particularly for pathogens transmitted in the hospital environment.

Recombinant antibodies against SlpA, FliC and FliD were also studied in electron microscopy using Protein A-gold but results were disappointing when compared to other studies. In most reports of *C. difficile* that use these methods, polyclonal antisera were used, raised against native or recombinant protein targets. The success of this approach is illustrated in studies of the presence of flagellae on strains of *C. difficile*. The binding of antibodies was visualised using colloidal gold particles as in the present investigation but with better outcome (Tasteyre et al., 2000a). Similarly, the presence of CwpV at the surface of *C. difficile* and its phase variation was confirmed *via* these methods (Emerson et al., 2009) and GroEL was

detected by both immunoelectron microscopy and indirect immunofluorescence (Hennequin et al., 2001b).

Further, the exposure of amino and carboxy-terminal domains of Cwp66 at the bacterial surface was assessed using antisera raised against different regions of Cwp66 and immunogold techniques (Waligora et al., 2001). Again, had time allowed technical problems to be defined and overcome, the availability of scFvs against defined parts of Cwp66 would have allowed surface location and exposure to be determined. In future investigations, one option worth exploring is the recloning of scFv sequences to fuse a metal binding domain at the carboxy-terminus. This would allow sensitive but direct detection of binding of the scFv to its target (Malecki et al., 2002).

4.4 Conclusion

Overall, this project has validated the strategy of generating recombinant antibodies against selected targets from a bacterial pathogen using genomic data as the starting point and accomplishing the aims set *via* heterologous expression of the chosen targets and isolation of specific scFv antibodies from a synthetically diversified library by phage display. The work has confirmed the Tomlinson humanised antibody libraries as an excellent source from which to select scFv binders against *C. difficile* targets. The scFvs generally carried unique CDR sequences suggesting binding sites for each target were different. All antibodies were checked for recognition of the recombinant target proteins and in many cases, the scFvs were able to bind the native proteins in extracts from *C. difficile*. In only one area was it possible to assess whether scFvs possessed the ability to block the action of the native protein but the study confirmed that anti-FliC and anti-FliD scFvs were able inhibit the motility of *C. difficile*.

The strategy was totally dependent on the ability to express and purify clostridial target proteins in a heterologous host, *E. coli* and while all eleven target sequences were recovered from the genome of *C. difficile* 630 by PCR and cloned into an expression vector, three (Cwp84, FbpA, Acd) could not be expressed. For the remaining eight plus purified LMW protein from SlpA, phage display successfully yielded specific scFv antibodies. For LMW, FliC that are known to be variable between strains of *C. difficile*, scFvs appeared to recognise strain-specific epitopes.

This was fortuitous as the phage display strategy was not steered towards this outcome. For some targets (*eg* CspA, sortase B), scFvs were reactive with the recombinant form of the protein but not with clostridial extracts. This might reflect abundance of the targets in the bacterial extracts. The study identified scFv-binding activity in extracts of *C. difficile* M120 and this is putatively attributed to

the LMW component of SlpA in this strain. Subject to the some limitations, the project showed that phage display can be used as a rapid method to generate specific reagents against a wide range of proteins from an important bacterial pathogen. These reagents have potential in diagnostics, in advancing understanding of the disease process, and in developing new therapeutics.

Immunofluorescence microscopy was used with scFvs against FliC and FliD and Protein A-FITC. Given the nature of the bacterial target, it was expected that this would generate a pattern of discrete labelling on the surface of the clostridial cells. Instead, uniform staining was observed over the entire surface of the cells. This was suggestive of non-specific staining. To have a definite pattern for those recombinant antibodies as a detached green signal from the body of bacteria this revealed a whole green exposure for all body.

This aspect of the project would have benefited from better choice of controls. For example, the scFv could have been omitted to assess if the FITC Protein-A conjugate was binding directly to the bacterial surface. Alternatively, recombinant antibodies could have been used in the staining protocol that were against targets not found on the surface of *C.difficile*, for instance anti-BSA, anti-ubiquitin or scFvs against surface proteins from other pathogens. These variations to the method could have helped clarify if in the pattern of immunofluorescence that was observed with anti-FliC and anti-FliD scFvs was specific.

The other approach to checking the immunofluorescence result would have been to use as target a knockout strain of *C. difficile* that was unable to synthesise flagellar subunits or assemble them into intact surface structures. For instance, inactivation of the glycosyltransferase gene present at the flagellar locus is known to lead to a failure to produce flagellae (Twine et al., 2009). Alternatively, naturally non-flagellated strains of *C. difficile* like M120 would make good controls though evidence collected here suggests that this particular strain may be able to capture scFvs through other processes. Overall, the immunofluorescence experiments

reported here show a staining pattern suggestive of uniform binding to the bacterial surface. Use of good controls could assess the reliability and reproducibility of the data reported.

It was unfortunate that limitations of time prevented a systematic investigation of why the binding of scFvs to the surface of *C. difficile* could not be detected in electron microscopy. With more time, investigation could have been carried out to assess whether sample preparation was a factor, specifically if fixation of the samples altered the clostridial targets in such a way that scFvs were unable to bind. Alternatively, the problem may have arisen in the interaction between scFvs and the Protein A-gold. Recent work has reported the direct conjugation of scFvs with gold nanoparticles. This interaction has been engineered by the introduction of cysteine or histidine residues into the scFv linker region and then coupling the protein with gold nanoparticles (Liu et al., 2009). In other work, metal binding domains have been fused to scFvs. This approach has been shown to create specific labelling with clear backgrounds and has been regarded by some researchers as more effective than conjugation with gold (Malecki et al., 2002).

While these aspects of the project left some issues unresolved, the ability of scFvs to inhibit bacterial motility offered a more satisfactory outcome. Results could have been enhanced with a more quantitative approach to the measurement of bacterial motility although this is complicated by the need to maintain the organisms under anaerobic conditions. Further enhancement could have come from a greater range of controls (*eg* non-motile strains created by gene knockouts) and the use of a wider range of motile *C. difficile* strains. Potentially, this could have been used to map the epitopes recognised by anti-FliC and anti-FliD scFvs.

Again, with more time other scFvs could have been tested for the capacity to block the function of their respective targets. *In vitro* assays exist for GroEL activity. For example, it was reported the interaction of oxidized GroEL with

unfolded proteins can be monitored at low temperature. Oxidized GroEL then can efficiently bind to non native substrates at low temperature with out ability of ATPase activity that can be reserved at higher temperature leading to release of native substrates (Melkani et al.) and in the absence of GroES, GroEL carries out potassium-dependant hydrolysis of ATP (Todd et al., 1993). These activities could be assayed in the absence and presence of anti-GroEL scFvs and, as controls, scFvs against other targets.

Other possibilities exist for future work. For example, Cwp66 has an ability to bind to cultured mammalian cells like Hep-2 or to gastrointestinal tissue. By inducing expression of Cwp66 through stress and checking its presence at the bacterial surface, there is potential here to assess the effect of all the selected recombinant antibodies on bacterial adherence *in vitro*. Given that scFvs were isolated against both the amino and carboxy terminal regions of Cwp66, these experiments could also clarify those regions of the protein that are responsible for its adherent activity.

5 Appendices

5.1 Brazier's CCEY agar

Brazier's powder (Oxoid) 48.0 g to 1 litre deionised water

The solution was autoclaved at 121 °C for 15 minutes. After cooling down to 47°C, 10 ml of cycloserine/cefoxitin(X093, LAB M) and 40 ml egg yolk emulsion (X073, LAB M) were added as supplement.

5.2 BHI liquid medium

BHI powder (Oxoid) 12.0 g

The powder was dissolved in 1 litre deionised water and autoclaved at 121°C for 15 minutes.

5.3 Agarose solution in TEA buffer

Agarose (invitrogene, Cat No. 15510-027) 1g

TEA buffer (Tris-EDTA-Acetate, Amresco) 4 ml of 25X

After adding the volume of buffer to 100 ml, agarose was dissolved in it and autoclaved at 121°C for 15 minutes. After cooling it down, 3 µl of SYBR safe (invitrogen) was added and poured to a gel casting tray.

5.4 2xYT medium

Tryptone 16g

Yeast extract 10g

Nacl 5g

Above materials dissolved to 1 litre deionised water. The broth was autoclaved at 121°C for 15 minutes. To prepare an agar base of this medium, 15g agar can be added to all above material, then the same conditions used to autoclaved.

5.5 Antibiotic solutions

Stock solutions of ampicillin (1g/ml) and kanamycin (0.5g/ml) were prepared in distilled water. Solutions were filtered by passing them through a filter with 0.2µm pore size and stored at -20°C.

5.6 Western blot solutions

5.6.1 *Electroblotting buffer*

Tris base	7.2 g
Glycine	33.4 g

Dissolve the above in 2L distilled water, add 600 ml methanol and make up to 3L with distilled water.

5.6.2 *Ponceau S solution*

Ponceau S	0.5 g
-----------	-------

Glacial acetic acid 1 ml. Make up to 100 ml with distilled water.

5.6.3 *Blocking buffer*

Marvel skimmed milk (2% w/v)	4 g
------------------------------	-----

Dissolved in 200ml PBS

References

- ABOUDOLA, S., KOTLOFF, K. L., KYNE, L., WARNY, M., KELLY, E. C., SOUGIOULTZIS, S., GIANNASCA, P. J., MONATH, T. P. & KELLY, C. P. (2003) Clostridium difficile vaccine and serum immunoglobulin G antibody response to toxin A. *Infection and Immunity*, 71, 1608-1610.
- ABOUGERGI, M. & KWON, J. Intravenous Immunoglobulin for the Treatment of Clostridium difficile Infection: A Review. *Digestive Diseases and Sciences*, 56, 19-26.
- AL SAIF, N. & BRAZIER, J. S. (1996) The distribution of Clostridium difficile in the environment of South Wales. *J Med Microbiol*, 45, 133-7.
- ALDEEN, W. E., BINGHAM, M., AIDERZADA, A., KUCERA, J., JENSE, S. & CARROLL, K. C. (2000) Comparison of the TOX A/B test to a cell culture cytotoxicity assay for the detection of Clostridium difficile in stools. *Diagnostic Microbiology and Infectious Disease*, 36, 211-213.
- AMERSDORFER, P., WONG, C., SMITH, T., CHEN, S., DESHPANDE, S., SHERIDAN, R. & MARKS, J. D. (2002) Genetic and immunological comparison of anti-botulinum type A antibodies from immune and non-immune human phage libraries. *Vaccine*, 20, 1640-1648.
- ARAP, M. A. (2005) Phage display technology - Applications and innovations. *Genetics and Molecular Biology*, 28, 1-9.
- ARAP, W., HAEDICKE, W., BERNASCONI, M., KAIN, R., RAJOTTE, D., KRAJEWSKI, S., ELLERBY, H. M., BREDESEN, D. E., PASQUALINI, R. & RUOSLAHTI, E. (2002) Targeting the prostate for destruction through a vascular address. *Proceedings of the National Academy of Sciences*, 99, 1527-1531.
- ARBABI-GHAHROUDI, M., TANHA, J. & MACKENZIE, R. (2005) Prokaryotic expression of antibodies. *Cancer and Metastasis Reviews*, 24, 501-519.
- ARBABI GHAHROUDI, M., DESMYTER, A., WYNS, L., HAMERS, R. & MUYLDERMANS, S. (1997) Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Letters*, 414, 521-526.
- ARROYO, L. G., KRUTH, S. A., WILLEY, B. M., STAEMPFLI, H. R., LOW, D. E. & WEESE, J. S. (2005) PCR ribotyping of Clostridium difficile isolates originating from human and animal sources. *Journal of Medical Microbiology*, 54, 163-166.
- AULABAUGH, A., DING, W., KAPOOR, B., TABEL, K., ALKSNE, L., DUSHIN, R., ZATZ, T., ELLESTAD, G. & HUANG, X. (2007) Development of an HPLC assay for Staphylococcus aureus sortase: Evidence for the formation of the kinetically competent acyl enzyme intermediate. *Analytical Biochemistry*, 360, 14-22.
- AUSIELLO, C. M., CERQUETTI, M., FEDELE, G., SPENSIERI, F., PALAZZO, R., NASSO, M., FREZZA, S. & MASTRANTONIO, P. (2006) Surface layer proteins from Clostridium difficile induce inflammatory and regulatory cytokines in human monocytes and dendritic cells. *Microbes Infect*, 8, 2640-6.
- BACCI, S., ST-MARTIN, G., OLESEN, B., BRUUN, B., OLSEN, K. E., NIELSEN, E. M. & MOLBAK, K. (2009) OUTBREAK OF CLOSTRIDIUM DIFFICILE 027 IN NORTH ZEALAND, DENMARK, 2008-2009. *Eurosurveillance*, 14, 4-6.
- BAEK, H., SUK, K. H., KIM, Y. H. & CHA, S. (2002) An improved helper phage system for efficient isolation of specific antibody molecules in phage display. *Nucleic Acids Research*, 30, 9.
- BAKRI, M. M., BROWN, D. J., BUTCHER, J. P. & SUTHERLAND, A. D. (2009) Clostridium difficile in Ready-to-Eat Salads, Scotland. *Emerging Infectious Diseases*, 15, 817-818.
- BARBAS, C. F., BAIN, J. D., HOEKSTRA, D. M. & LERNER, R. A. (1992) Semisynthetic combinatorial antibody libraries: a chemical solution to the diversity problem. *Proceedings of the National Academy of Sciences*, 89, 4457-4461.
- BARBAS, C. F., COLLET, T. A., AMBERG, W., ROBEN, P., BINLEY, J. M., HOEKSTRA, D., CABABA, D., JONES, T. M., WILLIAMSON, R. A., PILKINGTON, G. R., HAIGWOOD, N. L., CABEZAS, E., SATTERTHWAIT, A. C., SANZ, I. & BURTON, D. R. (1993) Molecular Profile of an Antibody Response to HIV-1 as Probed by Combinatorial Libraries. *Journal of Molecular Biology*, 230, 812-823.
- BARBAS, C. F., KANG, A. S., LERNER, R. A. & BENKOVIC, S. J. (1991) Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proceedings of the National Academy of Sciences*, 88, 7978-7982.
- BARBUT, F., DELMEE, M., BRAZIER, J. S., PETIT, J. C., POXTON, I. R., RUPNIK, M., LALANDE, V., SCHNEIDER, C., MASTRANTONIO, P., ALONSO, R., KUIPJER, E., TVEDE, M. & ESGCD (2003) A European survey of diagnostic methods and testing protocols for Clostridium difficile. *Clinical Microbiology and Infection*, 9, 989-996.

- BARBUT, F., MASTRANTONIO, P., DELMEE, M., BRAZIER, J., KUIJPER, E., POXTON, I. & EUROPEAN STUDY GRP CLOSTRIDIUM, D. (2007) Prospective study of *Clostridium difficile* infections in Europe with phenotypic and genotypic characterisation of the isolates. *Clinical Microbiology and Infection*, 13, 1048-1057.
- BARKER, R. H., DAGHER, R., DAVIDSON, D. M. & MARQUIS, J. K. (2006) Review article: tolevamer, a novel toxin-binding polymer: overview of preclinical pharmacology and physicochemical properties. *Alimentary Pharmacology & Therapeutics*, 24, 1525-1534.
- BARKETI-KLAI, A., HOYS, S., LAMBERT-BORDES, S., COLLIGNON, A. & KANSAU, I. Role of fibronectin-binding protein A in *Clostridium difficile* intestinal colonization. *J Med Microbiol*, 60, 1155-61.
- BARKETI-KLAI, A., HOYS, S., LAMBERT-BORDES, S., COLLIGNON, A. & KANSAU, I. Role of fibronectin binding protein A in *Clostridium difficile* intestinal colonization. *J Med Microbiol*, jmm.0.029553-0.
- BARNETT, T. C., PATEL, A. R. & SCOTT, J. R. (2004) A novel sortase, SrtC2, from *Streptococcus pyogenes* anchors a surface protein containing a QVPTGV motif to the cell wall. *J Bacteriol*, 186, 5865-75.
- BARNETT, T. C. & SCOTT, J. R. (2002) Differential recognition of surface proteins in *Streptococcus pyogenes* by two sortase gene homologs. *J Bacteriol*, 184, 2181-91.
- BARRY, M. A., DOWER, W. J. & JOHNSTON, S. A. (1996) Toward cell-targeting gene therapy vectors: Selection of cell-binding peptides from random peptide-presenting phage libraries. *Nature Medicine*, 2, 299-305.
- BARTH, H., PFEIFER, G., HOFMANN, F., MAIER, E., BENZ, R. & AKTORIES, K. (2001) Low pH-induced formation of ion channels by *Clostridium difficile* toxin B in target cells. *Journal of Biological Chemistry*, 276, 10670-10676.
- BARTLETT, J. G. (2002) Antibiotic-associated diarrhea. *New England Journal of Medicine*, 346, 334-339.
- BENHAR, I. & PASTAN, I. (1994) Cloning, expression and characterization of the Fv fragments of the anti-carbohydrate mAbs B1 and B5 as single-chain immunotoxins. *Protein Engineering*, 7, 1509-1515.
- BERGE, A. & BJORCK, L. (1995) Streptococcal cysteine proteinase releases biologically active fragments of streptococcal surface proteins. *J Biol Chem*, 270, 9862-7.
- BIERNE, H., MAZMANIAN, S. K., TROST, M., PUCCIARELLI, M. G., LIU, G., DEHOUX, P., JANSCH, L., GARCIA-DEL PORTILLO, F., SCHNEEWIND, O. & COSSART, P. (2002) Inactivation of the *srtA* gene in *Listeria monocytogenes* inhibits anchoring of surface proteins and affects virulence. *Mol Microbiol*, 43, 869-81.
- BOLKEN, T. C., FRANKE, C. A., JONES, K. F., ZELLER, G. O., JONES, C. H., DUTTON, E. K. & HRUBY, D. E. (2001) Inactivation of the *srtA* gene in *Streptococcus gordonii* inhibits cell wall anchoring of surface proteins and decreases in vitro and in vivo adhesion. *Infect Immun*, 69, 75-80.
- BOLTON, R. P., TAIT, S. K., DEAR, P. R. F. & LOSOWSKY, M. S. (1984) ASYMPTOMATIC NEONATAL COLONIZATION BY CLOSTRIDIUM-DIFFICILE. *Archives of Disease in Childhood*, 59, 466-472.
- BORGHESI, L. & MILCAREK, C. (2006) From B cell to plasma cell. *Immunologic Research*, 36, 27-32.
- BOURGAULT, A. M., LAMOTHE, F., LOO, V. G., POIRIER, L. & GRP, C.-C. S. (2006) In vitro susceptibility of *Clostridium difficile* clinical isolates from a multi-institutional outbreak in southern Quebec, Canada. *Antimicrobial Agents and Chemotherapy*, 50, 3473-3475.
- BRADBURY, A., PERSIC, L., WERGE, T. & CATTANEO, A. (1993) USE OF LIVING COLUMNS TO SELECT SPECIFIC PHAGE ANTIBODIES. *Bio-Technology*, 11, 1565-1569.
- BRAZIER, J. S. (2001) Typing of *Clostridium difficile*. *Clinical Microbiology and Infection*, 7, 428-431.
- BREKKE, O. H. & SANDLIE, I. (2003) Therapeutic antibodies for human diseases at the dawn of the twenty-first century. *Nat Rev Drug Discov*, 2, 52-62.
- BRILES, D. E., TART, R. C., SWIATLO, E., DILLARD, J. P., SMITH, P., BENTON, K. A., RALPH, B. A., BROOKS-WALTER, A., CRAIN, M. J., HOLLINGSHEAD, S. K. & MCDANIEL, L. S. (1998) Pneumococcal diversity: considerations for new vaccine strategies with emphasis on pneumococcal surface protein A (PspA). *Clin Microbiol Rev*, 11, 645-57.
- BRINKMANN, U., CHOWDHURY, P. S., ROSCOE, D. M. & PASTAN, I. (1995) Phage display of disulfide-stabilized Fv fragments. *Journal of Immunological Methods*, 182, 41-50.
- BRUGGEMANN, H. & GOTTSCHALK, G. (2008) Comparative genomics of clostridia - Link between the ecological niche and cell surface properties. *Incredible Anaerobes: From Physiology to Genomics to Fuels*, 1125, 73-81.
- BRUN, P., SCARPA, M., GRILLO, A., PALU, G., MENGOLI, C., ZECONI, A., SPIGAGLIA, P., MASTRANTONIO, P. & CASTAGLIUOLO, I. (2008) *Clostridium difficile* TxA(C314) and SLP-36kDa enhance the immune response toward a co-administered antigen. *Journal of Medical Microbiology*, 57, 725-731.
- CAI, X. & GAREN, A. (1995) Anti-melanoma antibodies from melanoma patients immunized with genetically modified autologous tumor cells: selection of specific antibodies from single-chain Fv fusion phage libraries. *Proceedings of the National Academy of Sciences*, 92, 6537-6541.

- CALABI, E., CALABI, F., PHILLIPS, A. D. & FAIRWEATHER, N. F. (2002) Binding of *Clostridium difficile* surface layer proteins to gastrointestinal tissues. *Infect Immun*, 70, 5770-8.
- CALABI, E. & FAIRWEATHER, N. (2002) Patterns of sequence conservation in the S-Layer proteins and related sequences in *Clostridium difficile*. *J Bacteriol*, 184, 3886-97.
- CALABI, E., WARD, S., WREN, B., PAXTON, T., PANICO, M., MORRIS, H., DELL, A., DOUGAN, G. & FAIRWEATHER, N. (2001) Molecular characterization of the surface layer proteins from *Clostridium difficile*. *Mol Microbiol*, 40, 1187-99.
- CARTER, P., KELLEY, R. F., RODRIGUES, M. L., SNEDECOR, B., COVARRUBIAS, M., VELLIGAN, M. D., WONG, W. L. T., ROWLAND, A. M., KOTTS, C. E., CARVER, M. E., YANG, M., BOURELL, J. H., SHEPARD, H. M. & HENNER, D. (1992) High Level *Escherichia coli* Expression and Production of a Bivalent Humanized Antibody Fragment. *Nat Biotech*, 10, 163-167.
- CASADEVALL, A. (2003) Antibody-Mediated Immunity against Intracellular Pathogens: Two-Dimensional Thinking Comes Full Circle. *Infect. Immun.*, 71, 4225-4228.
- CERQUETTI, M., MOLINARI, A., SEBASTIANELLI, A., DIOCIUTI, M., PETRUZZELLI, R., CAPO, C. & MASTRANTONIO, P. (2000a) Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microb Pathog*, 28, 363-72.
- CERQUETTI, M., MOLINARI, A., SEBASTIANELLI, A., DIOCIUTI, M., PETRUZZELLI, R., CAPO, C. & MASTRANTONIO, P. (2000b) Characterization of surface layer proteins from different *Clostridium difficile* clinical isolates. *Microbial Pathogenesis*, 28, 363-372.
- CERQUETTI, M., PANTOSTI, A., STEFANELLI, P. & MASTRANTONIO, P. (1992a) PURIFICATION AND CHARACTERIZATION OF AN IMMUNODOMINANT-36 KDA ANTIGEN PRESENT ON THE CELL-SURFACE OF CLOSTRIDIUM-DIFFICILE. *Microbial Pathogenesis*, 13, 271-279.
- CERQUETTI, M., PANTOSTI, A., STEFANELLI, P. & MASTRANTONIO, P. (1992b) Purification and characterization of an immunodominant 36 kDa antigen present on the cell surface of *Clostridium difficile*. *Microb Pathog*, 13, 271-9.
- CERQUETTI, M., SERAFINO, A., SEBASTIANELLI, A. & MASTRANTONIO, P. (2002) Binding of *Clostridium difficile* to Caco-2 epithelial cell line and to extracellular matrix proteins. *FEMS Immunol Med Microbiol*, 32, 211-8.
- CHANG, C., LANDOLFI, N. & QUEEN, C. (1991) Expression of antibody Fab domains on bacteriophage surfaces. Potential use for antibody selection. *The Journal of Immunology*, 147, 3610-3614.
- CHAPETON MONTES, D., CANDELA, T., COLLIGNON, A. & JANOIR, C. Localization of the *Clostridium difficile* cysteine protease Cwp84 and insights into its maturation process. *J Bacteriol*.
- CHAPETON MONTES, D., CANDELA, T., COLLIGNON, A. & JANOIR, C. Localization of the *Clostridium difficile* cysteine protease Cwp84 and insights into its maturation process. *J. Bacteriol.*, JB.00326-11.
- CHEN, Z., MOAYERI, M., ZHOU, Y.-H., LEPLA, S., EMERSON, S., SEBRELL, A., YU, F., SVITEL, J., SCHUCK, P., ST. CLAIRE, M. & PURCELL, R. (2006) Efficient Neutralization of Anthrax Toxin by Chimpanzee Monoclonal Antibodies against Protective Antigen. *Journal of Infectious Diseases*, 193, 625-633.
- CHENG, V. C. C., YAM, W. C., CHAN, J. F. W., TO, K. K. W., HO, P. L. & YUEN, K. Y. (2009) *Clostridium difficile* ribotype 027 arrives in Hong Kong. *International Journal of Antimicrobial Agents*, 34, 492-493.
- CHINTALACHARUVU, K. R. & MORRISON, S. L. (1995) Chimeric Antibodies: Production and Applications. *Methods*, 8, 73-82.
- CHOTHIA, C., LESK, A. M., TRAMONTANO, A., LEVITT, M., SMITH-GILL, S. J., AIR, G., SHERIFF, S., PADLAN, E. A., DAVIES, D., TULIP, W. R., COLMAN, P. M., SPINELLI, S., ALZARI, P. M. & POLJAK, R. J. (1989) Conformations of immunoglobulin hypervariable regions. *Nature*, 342, 877-883.
- CHOU, T.-H. W., WANG, S., SAKHATSKYY, P. V., MBOUDJEKA, I., LAWRENCE, J. M., HUANG, S., COLEY, S., YANG, B., LI, J., ZHU, Q. & LU, S. (2005) Epitope mapping and biological function analysis of antibodies produced by immunization of mice with an inactivated Chinese isolate of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) (vol 334, pg 134, 2005). *Virology*, 337, 204.
- CHRISTIAN, R. B., ZUCKERMANN, R. N., KERR, J. M., WANG, L. & MALCOLM, B. A. (1992) Simplified methods for construction, assessment and rapid screening of peptide libraries in bacteriophage. *Journal of Molecular Biology*, 227, 711-718.
- CLACKSON, T., HOOGENBOOM, H. R., GRIFFITHS, A. D. & WINTER, G. (1991) Making antibody fragments using phage display libraries. *Nature*, 352, 624-628.
- CLARK, J. R. & MARCH, J. B. (2004a) Bacterial viruses as human vaccines? *Expert Review of Vaccines*, 3, 463-476.
- CLARK, J. R. & MARCH, J. B. (2004b) Bacteriophage-mediated nucleic acid immunisation. *FEMS Immunology & Medical Microbiology*, 40, 21-26.

- CLARK, M. A., HAMMOND, F. R., PAPAIOANNOU, A., HAWKINS, N. J. & WARD, R. L. (1997) Regulation and expression of human Fabs under the control of the Escherichia coli arabinose promoter, P-BAD. *Immunotechnology*, 3, 217-226.
- COLCHER, D., BIRD, R., ROSELLI, M., HARDMAN, K. D., JOHNSON, S., POPE, S., DODD, S. W., PANTOLIANO, M. W., MILENIC, D. E. & SCHLOM, J. (1990) In Vivo Tumor Targeting of a Recombinant Single-Chain Antigen-Binding Protein. *Journal of the National Cancer Institute*, 82, 1191-1197.
- COOKE, D. L. & BORRIELLO, S. P. (1998) Nonspecific Binding of Clostridium difficile Toxin A to Murine Immunoglobulins Occurs via the Fab Component. *Infection and Immunity*, 66, 1981-1984.
- CROBACH, M. J. T., DEKKERS, O. M., WILCOX, M. H. & KUIJPER, E. J. (2009) European Society of Clinical Microbiology and Infectious Diseases (ESCMID): Data review and recommendations for diagnosing Clostridium difficile-infection (CDI). *Clinical Microbiology and Infection*, 15, 1053-1066.
- CUIV, P. O., KLAASSENS, E. S., DURKIN, A. S., HARKINS, D. M., FOSTER, L., MCCORRISON, J., TORRALBA, M., NELSON, K. E. & MORRISON, M. Draft Genome Sequence of Turicibacter sanguinis PC909, Isolated from Human Feces. *Journal of Bacteriology*, 193, 1288-1289.
- CUMMINGS, P. J., HOOPER, N. E. & ROWLAND, S. S. (1998) Generation of a recombinant bacteriophage antibody library to mycobacterium tuberculosis. *Hybridoma*, 17, 151-6.
- CWIRLA, S. E., PETERS, E. A., BARRETT, R. W. & DOWER, W. J. (1990) PEPTIDES ON PHAGE - A VAST LIBRARY OF PEPTIDES FOR IDENTIFYING LIGANDS. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 6378-6382.
- DALLAL, R. M., HARBRECHT, B. G., BOUJOUKAS, A. J., SIRIO, C. A., FARKAS, L. M., LEE, K. K. & SIMMONS, R. L. (2002) Fulminant Clostridium difficile: An underappreciated and increasing cause of death and complications. *Annals of Surgery*, 235, 363-372.
- DALLAS, S. D. & ROLFE, R. D. (1998) Binding of Clostridium difficile toxin A to human milk secretory component. *Journal of Medical Microbiology*, 47, 879-888.
- DALTON, B. R., LYE-MACCANNELL, T., HENDERSON, E. A., MACCANNELL, D. R. & LOUIE, T. J. (2009) Proton pump inhibitors increase significantly the risk of Clostridium difficile infection in a low-endemicity, non-outbreak hospital setting. *Aliment Pharmacol Ther*, 29, 626-34.
- DAVIES, E. L., SMITH, J. S., BIRKETT, C. R., MANSER, J. M., ANDERSON-DEAR, D. V. & YOUNG, J. R. (1995) Selection of specific phage-display antibodies using libraries derived from chicken immunoglobulin genes. *Journal of Immunological Methods*, 186, 125-135.
- DE LA RIVA, L., WILLING, S. E., TATE, E. W. & FAIRWEATHER, N. F. Roles of cysteine proteases Cwp84 and Cwp13 in biogenesis of the cell wall of Clostridium difficile. *J Bacteriol*, 193, 3276-85.
- DE WILDT, R. M., TOMLINSON, I. M., ONG, J. L. & HOLLIGER, P. (2002) Isolation of receptor-ligand pairs by capture of long-lived multivalent interaction complexes. *Proc Natl Acad Sci U S A*, 99, 8530-5.
- DE WILDT, R. M. T., FINNERN, R., OUWEHAND, W. H., GRIFFITHS, A. D., VAN VENROOIJ, W. J. & HOET, R. M. A. (1996) Characterization of human variable domain antibody fragments against the U1 RNA-associated A protein, selected from a synthetic and a patient-derived combinatorial V gene library. *European Journal of Immunology*, 26, 629-639.
- DEBABOV, V. G. (2004) Bacterial and archaeal S-layers as a subject of nanobiotechnology. *Molecular Biology*, 38, 482-493.
- DEBAST, S. B., LEENGOED, L. A. M. G. V., GOORHUIS, A., HARMANUS, C., KUIJPER, E. J. & BERGWERFF, A. A. (2009a) Clostridium difficile PCR ribotype 078 toxinotype V found in diarrhoeal pigs identical to isolates from affected humans. *Environmental Microbiology*, 11, 505-511.
- DEBAST, S. B., VAN LEENGOED, L., GOORHUIS, A., HARMANUS, C., KUIJPER, E. J. & BERGWERFF, A. A. (2009b) Clostridium difficile PCR ribotype 078 toxinotype V found in diarrhoeal pigs identical to isolates from affected humans. *Environmental Microbiology*, 11, 505-511.
- DELMEE, M., AVESANI, V., DELFERRIERE, N. & BURTONBOY, G. (1990) Characterization of flagella of Clostridium difficile and their role in serogrouping reactions. *J Clin Microbiol*, 28, 2210-4.
- DELMEE, M., HOMEL, M. & WAUTERS, G. (1985) Serogrouping of Clostridium difficile strains by slide agglutination. *J Clin Microbiol*, 21, 323-7.
- DENG, X. K., NESBIT, L. A. & MORROW, K. J., JR. (2003) Recombinant Single-Chain Variable Fragment Antibodies Directed against Clostridium difficile Toxin B Produced by Use of an Optimized Phage Display System. *Clin. Diagn. Lab. Immunol.*, 10, 587-595.
- DENNIS, M. S. & LAZARUS, R. A. (1994) Kunitz domain inhibitors of tissue factor-factor VIIa. II. Potent and specific inhibitors by competitive phage selection. *Journal of Biological Chemistry*, 269, 22137-22144.
- DESIDERIO, A., FRANCONI, R., LOPEZ, M., VILLANI, M. E., VITI, F., CHIARALUCE, R., CONSALVI, V., NERI, D. & BENVENUTO, E. (2001) A semi-synthetic repertoire of intrinsically stable antibody fragments derived from a single-framework scaffold. *Journal of Molecular Biology*, 310, 603-615.

- DEVLIN, J. J., PANGANIBAN, L. C. & DEVLIN, P. E. (1990) RANDOM PEPTIDE LIBRARIES - A SOURCE OF SPECIFIC PROTEIN-BINDING MOLECULES. *Science*, 249, 404-406.
- DHALLUIN, A., BOURGEOIS, I., PESTEL-CARON, M., CAMIADE, E., RAUX, G., COURTIN, P., CHAPOT-CHARTIER, M. P. & PONS, J. L. (2005) Acd, a peptidoglycan hydrolase of *Clostridium difficile* with N-acetylglucosaminidase activity. *Microbiology*, 151, 2343-51.
- DIAL, S., ALRASADI, K., MANOUKIAN, C., HUANG, A. & MENZIES, D. (2004) Risk of *Clostridium difficile* diarrhea among hospital inpatients prescribed proton pump inhibitors: cohort and case-control studies. *Canadian Medical Association Journal*, 171, 33-38.
- DIAL, S., DELANEY, J. A. C., BARKUN, A. N. & SUISSA, S. (2005) Use of gastric acid-suppressive agents and the risk of community-acquired *Clostridium difficile*-associated disease. *Jama-Journal of the American Medical Association*, 294, 2989-2995.
- DINGLE, T. C., MULVEY, G. L. & ARMSTRONG, G. D. Mutagenic Analysis of the *Clostridium difficile* Flagellar Proteins, FliC and FliD, and Their Contribution to Virulence in Hamsters. *Infect. Immun.*, 79, 4061-4067.
- DONTA, S. T. & MYERS, M. G. (1982) CLOSTRIDIUM DIFFICILE TOXIN IN ASYMPTOMATIC-NEONATES. *Journal of Pediatrics*, 100, 431-434.
- DOORBAR, J. & WINTER, G. (1994) Isolation of a Peptide Antagonist to the Thrombin Receptor using Phage Display. *Journal of Molecular Biology*, 244, 361-369.
- DORAN, J. D., NOMIZU, M., TAKEBE, S., MENARD, R., GRIFFITH, D. & ZIOMEK, E. (1999) Autocatalytic processing of the streptococcal cysteine protease zymogen: processing mechanism and characterization of the autoproteolytic cleavage sites. *Eur J Biochem*, 263, 145-51.
- DRUDY, D., CALABI, E., KYNE, L., SOUGIOLTZIS, S., KELLY, E., FAIRWEATHER, N. & KELLY, C. P. (2004) Human antibody response to surface layer proteins in *Clostridium difficile* infection. *FEMS Immunol Med Microbiol*, 41, 237-42.
- DUEÑAS, M., AYALA, M., VJZQUEZ, J., OHLIN, M., SDERLIND, E., BORREBAECK, C. A. K. & GAVILONDO, J. V. (1995) A point mutation in a murine immunoglobulin V-region strongly influences the antibody yield in *Escherichia coli*. *Gene*, 158, 61-66.
- DURET, S., BERHO, N., DANET, J.-L., GARNIER, M. & RENAUDIN, J. L. (2003) Spiralin Is Not Essential for Helicity, Motility, or Pathogenicity but Is Required for Efficient Transmission of *Spiroplasma citri* by Its Leafhopper Vector *Circulifer haematoceps*. *Applied and Environmental Microbiology*, 69, 6225-6234.
- EIDHIN, D. N., RYAN, A. W., DOYLE, R. M., WALSH, J. B. & KELLEHER, D. (2006) Sequence and phylogenetic analysis of the gene for surface layer protein, slpA, from 14 PCR ribotypes of *Clostridium difficile*. *J Med Microbiol*, 55, 69-83.
- ELLIOTT, B., CHANG, B. J., GOLLEDGE, C. L. & RILEY, T. V. (2007) *Clostridium difficile*-associated diarrhoea. *Intern Med J*, 37, 561-8.
- EMERSON, J. E., REYNOLDS, C. B., FAGAN, R. P., SHAW, H. A., GOULDING, D. & FAIRWEATHER, N. F. (2009) A novel genetic switch controls phase variable expression of CwpV, a *Clostridium difficile* cell wall protein. *Molecular Microbiology*, 74, 541-556.
- ETIENNE-TOUMELIN, I., SIRARD, J. C., DUFLOT, E., MOCK, M. & FOUET, A. (1995) Characterization of the *Bacillus anthracis* S-layer: cloning and sequencing of the structural gene. *J Bacteriol*, 177, 614-20.
- EVEILLARD, M., FOUREL, V., BARC, M. C., KERNEIS, S., COCONNIER, M. H., KARJALAINEN, T., BOURLIOUX, P. & SERVIN, A. L. (1993) Identification and characterization of adhesive factors of *Clostridium difficile* involved in adhesion to human colonic enterocyte-like Caco-2 and mucus-secreting HT29 cells in culture. *Mol Microbiol*, 7, 371-81.
- FAGAN, R. P., ALBESA-JOVE, D., QAZI, O., SVERGUN, D. I., BROWN, K. A. & FAIRWEATHER, N. F. (2009) Structural insights into the molecular organization of the S-layer from *Clostridium difficile*. *Molecular Microbiology*, 71, 1308-1322.
- FAGAN, R. P., JANOIR, C., COLLIGNON, A., MASTRANTONIO, P., POXTON, I. R. & FAIRWEATHER, N. F. A proposed nomenclature for cell wall proteins of *Clostridium difficile*. *J Med Microbiol*, 60, 1225-8.
- FAGAN, R. P., JANOIR, C., COLLIGNON, A., MASTRANTONIO, P., POXTON, I. R. & FAIRWEATHER, N. F. A proposed nomenclature for cell wall proteins of *Clostridium difficile*. *J Med Microbiol*, jmm.0.028472-0.
- FELICI, F., CASTAGNOLI, L., MUSACCHIO, A., JAPPELLI, R. & CESARENI, G. (1991) SELECTION OF ANTIBODY LIGANDS FROM A LARGE LIBRARY OF OLIGOPEPTIDES EXPRESSED ON A MULTIVALENT EXPOSITION VECTOR. *Journal of Molecular Biology*, 222, 301-310.
- FELICI, F., LUZZAGO, A., FOLGORI, A. & CORTESE, R. (1993) Mimicking of discontinuous epitopes by phage-displayed peptides, II. Selection of clones recognized by a protective monoclonal antibody against the *Bordetella pertussis* toxin from phage peptide libraries. *Gene*, 128, 21-27.

- FENNER, L., WIDMER, A. F., STRANDEN, A., CONZELMANN, M., GOORHUIS, A., HARMANUS, C., KUIJPER, E. J. & FREI, R. (2008) First cluster of clindamycin-resistant *Clostridium difficile* PCR ribotype 027 in Switzerland. *Clinical Microbiology and Infection*, 14, 514-515.
- FIERABRACCI, A., BIRO, P. A., YIANGOU, Y., MENNUNI, C., LUZZAGO, A., LUDVIGSSON, J., CORTESE, R. & BOTTAZZO, G. F. (1999) Osteopontin is an autoantigen of the somatostatin cells in human islets: identification by screening random peptide libraries with sera of patients with insulin-dependent diabetes mellitus. *Vaccine*, 18, 342-354.
- FOLGORI, A., TAFI, R., MEOLA, A., FELICI, F., GALFRE, G., CORTESE, R., MONACI, P. & NICOSIA, A. (1994) A general strategy to identify mimotopes of pathological antigens using only random peptide libraries and human sera. *Embo J*, 13, 2236-43.
- FORDTRAN, J. S. (2006) Colitis due to *Clostridium difficile* toxins: underdiagnosed, highly virulent, and nosocomial. *Proc (Bayl Univ Med Cent)*, 19, 3-12.
- FORSBERG, G. R., FORSGREN, M., JAKI, M., NORIN, M., STERKY, C., ENHÄRNING, I. S., LARSSON, K., ERICSSON, M. & BJÖRK, P. (1997) Identification of Framework Residues in a Secreted Recombinant Antibody Fragment That Control Production Level and Localization in *Escherichia coli*. *Journal of Biological Chemistry*, 272, 12430-12436.
- FOWLKES, D. M., ADAMS, M. D., FOWLER, V. A. & KAY, B. K. (1992) MULTIPURPOSE VECTORS FOR PEPTIDE EXPRESSION ON THE M13 VIRAL SURFACE. *Biotechniques*, 13, 422-&.
- FREEMAN, J., BAINES, S. D., JABES, D. & WILCOX, M. H. (2005) Comparison of the efficacy of ramoplanin and vancomycin in both in vitro and in vivo models of clindamycin-induced *Clostridium difficile* infection. *Journal of Antimicrobial Chemotherapy*, 56, 717-725.
- FREEMAN, J., BAUER, M. P., BAINES, S. D., CORVER, J., FAWLEY, W. N., GOORHUIS, B., KUIJPER, E. J. & WILCOX, M. H. The Changing Epidemiology of *Clostridium difficile* Infections. *Clinical Microbiology Reviews*, 23, 529-549.
- GAILLARD, J. L., BERCHE, P., FREHEL, C., GOUIN, E. & COSSART, P. (1991) Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. *Cell*, 65, 1127-41.
- GAO, J., WANG, Y., LIU, Z. & WANG, Z. Phage display and its application in vaccine design. *Annals of Microbiology*, 60, 13-19.
- GASPAR, A. H., MARRAFFINI, L. A., GLASS, E. M., DEBORD, K. L., TON-THAT, H. & SCHNEEWIND, O. (2005) *Bacillus anthracis* sortase A (SrtA) anchors LPXTG motif-containing surface proteins to the cell wall envelope. *J Bacteriol*, 187, 4646-55.
- GAYNES, R., RIMLAND, D., KILLUM, E., LOWERY, H. K., JOHNSON, T. M., KILLGORE, G. & TENOVER, F. C. (2004) Outbreak of *Clostridium difficile* infection in a long-term care facility: Association with gatifloxacin use. *Clinical Infectious Diseases*, 38, 640-645.
- GERDING, D. N. (2005) Metronidazole for *Clostridium difficile*-associated disease: Is it okay for mom? *Clinical Infectious Diseases*, 40, 1598-1600.
- GIANNASCA, P. J. & WARNY, M. (2004a) Active and passive immunization against *Clostridium difficile* diarrhea and colitis. *Vaccine*, 22, 848-56.
- GIANNASCA, P. J. & WARNY, M. (2004b) Active and passive immunization against *Clostridium difficile* diarrhea and colitis (Reprinted from 22(7)848 Feb 17 2004). *Vaccine*, 22, 848-856.
- GIBBS, R. A., POSNER, B. A., FILPULA, D. R., DODD, S. W., FINKELMAN, M. A., LEE, T. K., WROBLE, M., WHITLOW, M. & BENKOVIC, S. J. (1991) Construction and characterization of a single-chain catalytic antibody. *Proceedings of the National Academy of Sciences*, 88, 4001-4004.
- GOLDENBERG, S. D., CLIFF, P. R. & FRENCH, G. L. Glutamate Dehydrogenase for Laboratory Diagnosis of *Clostridium difficile* Infection. *Journal of Clinical Microbiology*, 48, 3050-3051.
- GOLDMAN, C. K., ROGERS, B. E., DOUGLAS, J. T., SOSNOWSKI, B. A., YING, W., SIEGAL, G. P., BAIRD, A., CAMPAIN, J. A. & CURIEL, D. T. (1997) Targeted Gene Delivery to Kaposi's Sarcoma Cells via the Fibroblast Growth Factor Receptor. *Cancer Research*, 57, 1447-1451.
- GOMEZTREVINO, M., BOUREAU, H., KARJALAINEN, T. & BOURLIOUX, P. (1996) *Clostridium difficile* adherence to mucus: Results of an in vivo and ex vivo assay. *Microbial Ecology in Health and Disease*, 9, 329-334.
- GOODSON, R. J., DOYLE, M. V., KAUFMAN, S. E. & ROSENBERG, S. (1994) High-affinity urokinase receptor antagonists identified with bacteriophage peptide display. *Proceedings of the National Academy of Sciences*, 91, 7129-7133.
- GREEN, L. L. (1999) Antibody engineering via genetic engineering of the mouse: XenoMouse strains are a vehicle for the facile generation of therapeutic human monoclonal antibodies. *Journal of Immunological Methods*, 231, 11-23.

- GREENWOOD, J., WILLIS, A. E. & PERHAM, R. N. (1991) Multiple display of foreign peptides on a filamentous bacteriophage : Peptides from Plasmodium falciparum circumsporozoite protein as antigens. *Journal of Molecular Biology*, 220, 821-827.
- GRIFFITHS, A. D., MALMQVIST, M., MARKS, J. D., BYE, J. M., EMBLETON, M. J., MCCAFFERTY, J., BAIER, M., HOLLIGER, K. P., GORICK, B. D. & HUGHES-JONES, N. C. (1993) Human anti-self antibodies with high specificity from phage display libraries. *The EMBO journal*, 12, 725-34.
- GRIFFITHS, A. D., WILLIAMS, S. C., HARTLEY, O., TOMLINSON, I. M., WATERHOUSE, P., CROSBY, W. L., KONTERMANN, R. E., JONES, P. T., LOW, N. M., ALLISON, T. J. & ET AL. (1994) Isolation of high affinity human antibodies directly from large synthetic repertoires. *Embo J*, 13, 3245-60.
- HAFIZ, S., MORTON, R. S., MCENTEGART, M. G. & WAITKINS, S. A. (1975) CLOSTRIDIUM DIFFICILE IN THE UROGENITAL TRACT OF MALES AND FEMALES. *The Lancet*, 305, 420-421.
- HALL I, O. T. E. (1935) International flora in newborn infants with a description of a new pathogenic anaerobe *Am J Dis Child* 49, 390
- HARRIS, B. (1999) Exploiting antibody-based technologies to manage environmental pollution. *Trends in Biotechnology*, 17, 290-296.
- HEAP, J. T., PENNINGTON, O. J., CARTMAN, S. T., CARTER, G. P. & MINTON, N. P. (2007) The Clostron: A universal gene knock-out system for the genus Clostridium. *Journal of Microbiological Methods*, 70, 452-464.
- HENNEQUIN, C., COLLIGNON, A. & KARJALAINEN, T. (2001a) Analysis of expression of GroEL (Hsp60) of Clostridium difficile in response to stress. *Microbial Pathogenesis*, 31, 255-260.
- HENNEQUIN, C., JANOIR, C., BARC, M. C., COLLIGNON, A. & KARJALAINEN, T. (2003) Identification and characterization of a fibronectin-binding protein from Clostridium difficile. *Microbiology*, 149, 2779-87.
- HENNEQUIN, C., PORCHERAY, F., WALIGORA-DUPRIET, A., COLLIGNON, A., BARC, M., BOURLIOUX, P. & KARJALAINEN, T. (2001b) GroEL (Hsp60) of Clostridium difficile is involved in cell adherence. *Microbiology*, 147, 87-96.
- HILL, D. F. & PETERSEN, G. B. (1982) NUCLEOTIDE-SEQUENCE OF BACTERIOPHAGE-F1 DNA. *Journal of Virology*, 44, 32-46.
- HO, J. G. S., GRECO, A., RUPNIK, M. & NG, K. K. S. (2005) Crystal structure of receptor-binding C-terminal repeats from Clostridium difficile toxin A. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 18373-18378.
- HOCHULI, E., BANNWARTH, W., DOBELI, H., GENTZ, R. & STUBER, D. (1988) GENETIC APPROACH TO FACILITATE PURIFICATION OF RECOMBINANT PROTEINS WITH A NOVEL METAL CHELATE ADSORBENT. *Bio-Technology*, 6, 1321-1325.
- HOLLIGER, P., RIECHMANN, L. & WILLIAMS, R. L. (1999) Crystal structure of the two N-terminal domains of g3p from filamentous phage fd at 1.9 angstrom: Evidence for conformational lability. *Journal of Molecular Biology*, 288, 649-657.
- HOLST, E., HELIN, I. & MARDH, P. A. (1981) RECOVERY OF CLOSTRIDIUM DIFFICILE FROM CHILDREN. *Scandinavian Journal of Infectious Diseases*, 13, 41-45.
- HOLTJE, J. V. & TOMASZ, A. (1975) Specific recognition of choline residues in the cell wall teichoic acid by the N-acetylmuramyl-L-alanine amidase of Pneumococcus. *J Biol Chem*, 250, 6072-6.
- HOOGENBOOM, H. R., GRIFFITHS, A. D., JOHNSON, K. S., CHISWELL, D. J., HUDSON, P. & WINTER, G. (1991) MULTISUBUNIT PROTEINS ON THE SURFACE OF FILAMENTOUS PHAGE - METHODOLOGIES FOR DISPLAYING ANTIBODY (FAB) HEAVY AND LIGHT-CHAINS. *Nucleic Acids Research*, 19, 4133-4137.
- HOOGENBOOM, H. R. & WINTER, G. (1992) By-passing immunisation : Human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro. *Journal of Molecular Biology*, 227, 381-388.
- HUSSACK, G. & TANHA, J. Toxin-Specific Antibodies for the Treatment of Clostridium difficile: Current Status and Future Perspectives. *Toxins*, 2, 998-1018.
- ILAN, E., BURAKOVA, T., DAGAN, S., NUSSBAUM, O., LUBIN, I., EREN, R., BEN-MOSHE, O., ARAZI, J., BERR, S., NEVILLE, L., YUEN, L., MANSOUR, T. S., GILLARD, J., EID, A., JURIM, O., SHOUVAL, D., REISNER, Y. & GALUN, E. (1999) The hepatitis B virus-trimer mouse: A model for human HBV infection and evaluation of Anti-HBV therapeutic agents. *Hepatology*, 29, 553-562.
- INDRA, A., HUHULESCU, S., FIEDLER, A., KERNBICHLER, S., BLASCHITZ, M. & ALLERBERGER, F. (2009) OUTBREAK OF CLOSTRIDIUM DIFFICILE 027 INFECTION IN VIENNA, AUSTRIA 2008-2009. *Eurosurveillance*, 14, 7-8.
- INDRA, A., HUHULESCU, S., HASENBERGER, P., SCHMID, D., ALFERY, C., WUERZNER, R., FILLE, M., GATTRINGER, K., KUIJPER, E. & ALLERBERGER, F. (2006) First isolation of Clostridium difficile PCR ribotype 027 in Austria. *Euro Surveill*, 11, E060914.3.

- INGEBRETSEN, A., HANSEN, G., HARMANUS, C. & KUIJPER, E. J. (2008) First confirmed cases of *Clostridium difficile* PCR ribotype 027 in Norway. *Euro Surveill*, 13.
- IRVING, M. B., PAN, O. & SCOTT, J. K. (2001) Random-peptide libraries and antigen-fragment libraries for epitope mapping and the development of vaccines and diagnostics. *Current Opinion in Chemical Biology*, 5, 314-324.
- JANK, T., GIESEMANN, T. & AKTORIES, K. (2007) Rho-glucosylating *Clostridium difficile* toxins A and B: new insights into structure and function. *Glycobiology*, 17, 15R-22R.
- JANOIR, C., PECHINE, S., GROSDIDIER, C. & COLLIGNON, A. (2007) Cwp84, a Surface-Associated Protein of *Clostridium difficile*, Is a Cysteine Protease with Degrading Activity on Extracellular Matrix Proteins. *J Bacteriol*, 189, 7174-80.
- JIANG, Z. D., DUPONT, H. L., GAREY, K., PRICE, M., GRAHAM, G., OKHUYSEN, P., DAO-TRAN, T. & LAROCCO, M. (2006) A common polymorphism in the interleukin 8 gene promoter is associated with *Clostridium difficile* diarrhea. *American Journal of Gastroenterology*, 101, 1112-1116.
- JIRHOLT, P., OHLIN, M., BORREBAECK, C. A. K. & SDERLIND, E. (1998) Exploiting sequence space: shuffling in vivo formed complementarity determining regions into a master framework. *Gene*, 215, 471-476.
- JOHANSEN, L. K., ALBRECHTSEN, B., ANDERSEN, H. W. & ENGBERG, J. (1995) pFab60: a new, efficient vector for expression of antibody Fab fragments displayed on phage. *Protein Engineering*, 8, 1063-1067.
- JOHNSON, S., SAMORE, M. H., FARROW, K. A., KILLGORE, G. E., TENOVER, F. C., LYRAS, D., ROOD, J. I., DEGIROLAMI, P., BALTCH, A. L., RAFFERTY, M. E., PEAR, S. M. & GERDING, D. N. (1999) Epidemics of diarrhea caused by a clindamycin-resistant strain of *Clostridium difficile* in four hospitals. *New England Journal of Medicine*, 341, 1645-1651.
- JONES, P. T., DEAR, P. H., FOOTE, J., NEUBERGER, M. S. & WINTER, G. (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. *Nature*, 321, 522-525.
- KANG, A. S., BARBAS, C. F., JANDA, K. D., BENKOVIC, S. J. & LERNER, R. A. (1991) Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. *Proceedings of the National Academy of Sciences*, 88, 4363-4366.
- KANITPUN, R., WAGNER, G. G. & WAGHELA, S. D. (2004) Characterization of recombinant antibodies developed for capturing enterohemorrhagic *Escherichia coli* O157:H7. *Southeast Asian J Trop Med Public Health*, 35, 902-12.
- KARJALAINEN, T., BARC, M. C., COLLIGNON, A., TROLLE, S., BOUREAU, H., COTTE-LAFFITTE, J. & BOURLIOUX, P. (1994) Cloning of a genetic determinant from *Clostridium difficile* involved in adherence to tissue culture cells and mucus. *Infect Immun*, 62, 4347-55.
- KARJALAINEN, T., SAUMIER, N., BARC, M. C., DELMEE, M. & COLLIGNON, A. (2002) *Clostridium difficile* genotyping based on slpA variable region in S-layer gene sequence: an alternative to serotyping. *J Clin Microbiol*, 40, 2452-8.
- KARJALAINEN, T., WALIGORA-DUPRIET, A. J., CERQUETTI, M., SPIGAGLIA, P., MAGGIONI, A., MAURI, P. & MASTRANTONIO, P. (2001) Molecular and genomic analysis of genes encoding surface-anchored proteins from *Clostridium difficile*. *Infect Immun*, 69, 3442-6.
- KARLE, S., PLANQUE, S., NISHIYAMA, Y., TAGUCHI, H., ZHOU, Y. X., SALAS, M., LAKE, D., THIAGARAJAN, P., ARNETT, F., HANSON, C. V. & PAUL, S. (2004) Cross-clade HIV-1 neutralization by an antibody fragment from a lupus phage display library. *AIDS*, 18, 329-31.
- KARLSSON, F., BORREBAECK, C. A. K., NILSSON, N. & MALMBORG-HAGER, A. C. (2003) The mechanism of bacterial infection by filamentous phages involves molecular interactions between TolA and phage protein 3 domains. *Journal of Bacteriology*, 185, 2628-2634.
- KATO, H., ITO, Y., AKAHANE, T., IZUMIDA, S., YOKOYAMA, T., KAJI, C. & ARAKAWA, Y. Typing of *Clostridium difficile* isolates endemic in Japan by sequencing of slpA and its application to direct typing. *Journal of Medical Microbiology*, 59, 556-562.
- KATO, H., ITO, Y., VAN DEN BERG, R. J., KUIJPER, E. J. & ARAKAWA, Y. (2007) First isolation of *Clostridium difficile* 027 in Japan. *Euro Surveill*, 12, E070111.3.
- KATO, H., KATO, H., ITO, Y., AKAHANE, T., IZUMIDA, S., YOKOYAMA, T., KAJI, C. & ARAKAWA, Y. Typing of *Clostridium difficile* isolates endemic in Japan by sequencing of slpA and its application to direct typing. *J Med Microbiol*, 59, 556-562.
- KATO, H., KATO, H., NAKAMURA, M., IWASHIMA, Y., NAKAMURA, A. & UEDA, R. (2009) Rapid analysis of *Clostridium difficile* strains recovered from hospitalized patients by using the slpA sequence typing system. *Journal of Infection and Chemotherapy*, 15, 199-202.
- KATO, H., KATO, N., WATANABE, K., YAMAMOTO, T., SUZUKI, K., ISHIGO, S., KUNIHIRO, S., NAKAMURA, I., KILLGORE, G. E. & NAKAMURA, S. (2001) Analysis of *Clostridium difficile* isolates from nosocomial outbreaks at three hospitals in diverse areas of Japan. *J Clin Microbiol*, 39, 1391-5.

- KATO, H., YOKOYAMA, T. & ARAKAWA, Y. (2005) Typing by sequencing the *slpA* gene of *Clostridium difficile* strains causing multiple outbreaks in Japan. *J Med Microbiol*, 54, 167-71.
- KAWATA, T., TAKEOKA, A., TAKUMI, K. & MASUDA, K. (1984) DEMONSTRATION AND PRELIMINARY CHARACTERIZATION OF A REGULAR ARRAY IN THE CELL-WALL OF CLOSTRIDIUM-DIFFICILE. *Fems Microbiology Letters*, 24, 323-328.
- KEEL, K., BRAZIER, J. S., POST, K. W., WEESE, S. & SONGER, J. G. (2007) Prevalence of PCR ribotypes among *Clostridium difficile* isolates from pigs, calves, and other species. *Journal of Clinical Microbiology*, 45, 1963-1964.
- KELLY, C. P., POTHOUAKIS, C. & LAMONT, J. T. (1994) CLOSTRIDIUM-DIFFICILE COLITIS - REPLY. *New England Journal of Medicine*, 330, 1755-1755.
- KETTLEBOROUGH, C. A., ANSELL, K. H., ALLEN, R. W., ROSELL-VIVES, E., GÜSSOW, D. H. & BENDIG, M. M. (1994) Isolation of tumor cell-specific single-chain Fv from immunized mice using phage-antibody libraries and the re-construction of whole antibodies from these antibody fragments. *European Journal of Immunology*, 24, 952-958.
- KIM, P.-H. & ROLFE, R. D. (1989) Characterisation of Protective Antibodies in Hamsters Immunised Against *Clostridium difficile* Toxins A and B. *Microbial Ecology in Health and Disease*, 2, 47-59.
- KIM, P. H., IACONIS, J. P. & ROLFE, R. D. (1987) IMMUNIZATION OF ADULT HAMSTERS AGAINST CLOSTRIDIUM-DIFFICILE-ASSOCIATED ILEOCEITIS AND TRANSFER OF PROTECTION TO INFANT HAMSTERS. *Infection and Immunity*, 55, 2984-2992.
- KIM, Y. J., LEBRETON, F., KAISER, C., CRUCIERE, C. & REMOND, M. (2004) Isolation of foot-and-mouth disease virus specific bovine antibody fragments from phage display libraries. *J Immunol Methods*, 286, 155-66.
- KIPRIYANOV, S. & LITTLE, M. (1999) Generation of recombinant antibodies. *Molecular Biotechnology*, 12, 173-201.
- KIPRIYANOV, S. M., MOLDENHAUER, G. & LITTLE, M. (1997) High level production of soluble single chain antibodies in small-scale *Escherichia coli* cultures. *Journal of Immunological Methods*, 200, 69-77.
- KIRBY, J. M., AHERN, H., ROBERTS, A. K., KUMAR, V., FREEMAN, Z., ACHARYA, K. R. & SHONE, C. C. (2009) Cwp84, a Surface-associated Cysteine Protease, Plays a Role in the Maturation of the Surface Layer of *Clostridium difficile*. *Journal of Biological Chemistry*, 284, 34666-34673.
- KIRBY, J. M., THIYAGARAJAN, N., ROBERTS, A. K., SHONE, C. C. & ACHARYA, K. R. Expression, purification, crystallization and preliminary crystallographic analysis of a putative *Clostridium difficile* surface protein Cwp19. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 67, 762-7.
- KNAPPIK, A., GE, L., HONEGGER, A., PACK, P., FISCHER, M., WELLNHOFER, G., HOESS, A., WILLE, J., PLÜCKTHUN, A. & VIRNEKÜS, B. (2000) Fully synthetic human combinatorial antibody libraries (HuCAL) based on modular consensus frameworks and CDRs randomized with trinucleotides. *Journal of Molecular Biology*, 296, 57-86.
- KOHLER, G. & MILSTEIN, C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256, 495-497.
- KRAMER, R. A., COX, F., VAN DER HORST, M., VAN DEN OUDENRIJN, S., RES, P. C. M., BIA, J., LOGTENBERG, T. & DE KRUIJF, J. (2003) A novel helper phage that improves phage display selection efficiency by preventing the amplification of phages without recombinant protein. *Nucleic Acids Research*, 31, 9.
- KRISTENSEN, P. & WINTER, G. (1998) Proteolytic selection for protein folding using filamentous bacteriophages. *Folding and Design*, 3, 321-328.
- KROOK, M., MOSBACH, K. & LINDBLADH, C. (1994) Selection of Peptides with Affinity for Single-Stranded DNA Using a Phage Display Library. *Biochemical and Biophysical Research Communications*, 204, 849-854.
- KUIJPER, E. J., VAN DEN BERG, R. J. & BRAZIER, J. S. (2009) Comparison of Molecular Typing Methods Applied to *Clostridium difficile*. IN CAUGANT, D. A. (Ed.) *Molecular Epidemiology of Microorganisms*. Humana Press Inc.
- KUIJPER, E. J., VAN DEN BERG, R. J., DEBAST, S., VISSER, C. E., VEENENDAAL, D., TROELSTRA, A., VAN DER KOOI, T., VAN DEN HOF, S. & NOTERMANS, D. W. (2006) *Clostridium difficile* ribotype 027, toxinotype III, the Netherlands. *Emerging Infectious Diseases*, 12, 827-830.
- LAFORCE-NESBITT, S. S., SULLIVAN, M. A., HOYER, L. L. & BLISS, J. M. (2008) Inhibition of *Candida albicans* adhesion by recombinant human antibody single-chain variable fragment specific for Als3p. *Fems Immunology and Medical Microbiology*, 54, 195-202.
- LAH, M., GOLDSTRAW, A., WHITE, J. F., DOLEZAL, O., MALBY, R. & HUDSON, P. J. (1994) Phage surface presentation and secretion of antibody fragments using an adaptable phagemid vector. *Human Antibodies and Hybridomas*, 5, 48-56.

- LAM, K. S., LEBL, M., KRCHNAK, V., WADE, S., ABDUL-LATIF, F., FERGUSON, R., CUZZOCREA, C. & WERTMAN, K. (1993) Discovery of d-amino-acid-containing ligands with selectide technology. *Gene*, 137, 13-16.
- LAMBRIANIDES, A. & GILES, I. Use of monoclonal antibodies to dissect specificity and pathogenesis of antiphospholipid antibodies. *Lupus*, 19, 359-364.
- LAROCCA, D., KASSNER, P. D., WITTE, A., LADNER, R. C., PIERCE, G. F. & BAIRD, A. (1999) Gene transfer to mammalian cells using genetically targeted filamentous bacteriophage. *The FASEB Journal*, 13, 727-734.
- LARRICK, J. W., DANIELSSON, L., BRENNER, C. A., ABRAHAMSON, M., FRY, K. E. & BORREBAECK, C. A. K. (1989) Rapid cloning of rearranged immunoglobulin genes from human hybridoma cells using mixed primers and the polymerase chain reaction. *Biochemical and Biophysical Research Communications*, 160, 1250-1256.
- LARSON, H. E., BARCLAY, F. E., HONOUR, P. & HILL, I. D. (1982) EPIDEMIOLOGY OF CLOSTRIDIUM-DIFFICILE IN INFANTS. *Journal of Infectious Diseases*, 146, 727-733.
- LARSON, H. E., PRICE, A. B., HONOUR, P. & BORRIELLO, S. P. (1978) CLOSTRIDIUM DIFFICILE AND ETIOLOGY OF PSEUDOMEMBRANOUS COLITIS. *Lancet*, 1, 1063-1066.
- LAUER, P., RINAUDO, C. D., SORIANI, M., MARGARIT, I., MAIONE, D., ROSINI, R., TADDEI, A. R., MORA, M., RAPPUOLI, R., GRANDI, G. & TELFORD, J. L. (2005) Genome analysis reveals pili in Group B Streptococcus. *Science*, 309, 105.
- LEE, J.-H., RHO, J. B., PARK, K.-J., KIM, C. B., HAN, Y.-S., CHOI, S. H., LEE, K.-H. & PARK, S.-J. (2004) Role of Flagellum and Motility in Pathogenesis of *Vibrio vulnificus*. *Infection and Immunity*, 72, 4905-4910.
- LEWIS, S., BURMEISTER, S. & BRAZIER, J. (2005) Effect of the prebiotic oligofructose on relapse of Clostridium difficile-associated diarrhea: A randomized, controlled study. *Clinical Gastroenterology and Hepatology*, 3, 442-448.
- LI, R., HOESS, R. H., BENNETT, J. S. & DEGRADO, W. F. (2003) Use of phage display to probe the evolution of binding specificity and affinity in integrins. *Protein Eng*, 16, 65-72.
- LIN, Y.-P., KUO, C.-J., KOLECI, X., MCDONOUGH, S. P. & CHANG, Y.-F. Manganese Binds to Clostridium difficile Fbp68 and Is Essential for Fibronectin Binding. *Journal of Biological Chemistry*, 286, 3957-3969.
- LIN, Y. P., KUO, C. J., KOLECI, X., MCDONOUGH, S. P. & CHANG, Y. F. Manganese binds to Clostridium difficile Fbp68 and is essential for fibronectin binding. *J Biol Chem*, 286, 3957-69.
- LITMAN, G. W., RAST, J. P., SHAMBLOTT, M. J., HAIRE, R. N., HULST, M., ROESS, W., LITMAN, R. T., HINDSFREY, K. R., ZILCH, A. & AMEMIYA, C. T. (1993) Phylogenetic diversification of immunoglobulin genes and the antibody repertoire. *Molecular Biology and Evolution*, 10, 60-72.
- LIU, Y., LIU, Y., MERNAUGH, R. L. & ZENG, X. (2009) Single chain fragment variable recombinant antibody functionalized gold nanoparticles for a highly sensitive colorimetric immunoassay. *Biosensors and Bioelectronics*, 24, 2853-2857.
- LOGAN, S. M. (2006) Flagellar glycosylation - a new component of the motility repertoire? *Microbiology-Sgm*, 152, 1249-1262.
- LOO, V. G. (2006) A predominantly clonal multi-institutional outbreak of Clostridium difficile-associated diarrhea with high morbidity and mortality (vol 353, pg 2442, 2005). *New England Journal of Medicine*, 354, 2200-2200.
- LORIMER, I. A. Å. J., KEPPLER-HAFKEMEYER, A., BEERS, R. A., PEGRAM, C. N., BIGNER, D. D. & PASTAN, I. (1996) Recombinant immunotoxins specific for a mutant epidermal growth factor receptor: Targeting with a single chain antibody variable domain isolated by phage display. *Proceedings of the National Academy of Sciences*, 93, 14815-14820.
- LOWMAN, H. B. & WELLS, J. A. (1993) Affinity Maturation of Human Growth Hormone by Monovalent Phage Display. *Journal of Molecular Biology*, 234, 564-578.
- LUPO, L., PANZERA, P., TANDOI, F., CARBOTTA, G., GIANNELLI, G., SANTANTONIO, T., RENDINA, M., GENTILE, A. & MEMEO, V. (2008) Basiliximab Versus Steroids in Double Therapy Immunosuppression in Liver Transplantation: A Prospective Randomized Clinical Trial. *Transplantation*, 86, 925-931 10.1097/TP.0b013e318186b8a3.
- LYERLY, D. M., BOSTWICK, E. F., BINION, S. B. & WILKINS, T. D. (1991) PASSIVE-IMMUNIZATION OF HAMSTERS AGAINST DISEASE CAUSED BY CLOSTRIDIUM-DIFFICILE BY USE OF BOVINE IMMUNOGLOBULIN-G CONCENTRATE. *Infection and Immunity*, 59, 2215-2218.
- LYRAS, D., O'CONNOR, J. R., HOWARTH, P. M., SAMBOL, S. P., CARTER, G. P., PHUMOONNA, T., POON, R., ADAMS, V., VEDANTAM, G., JOHNSON, S., GERDING, D. N. & ROOD, J. I. (2009) Toxin B is essential for virulence of Clostridium difficile. *Nature*, 458, 1176-1181.

- LYYTIKAINEN, O., MENTULA, S., KONONEN, E., KOTILA, S., TARKKA, E., ANTTILA, V. J., MATTILA, E., KANERVA, M., VAARA, M. & VALTONEN, V. (2007) First isolation of *Clostridium difficile* PCR ribotype 027 in Finland. *Euro Surveill*, 12, E071108.2.
- LYYTIKAINEN, O., TURUNEN, H., SUND, R., RASINPERA, M., KONONEN, E., RUUTU, P. & KESKIMAKI, I. (2009) Hospitalizations and Deaths Associated with *Clostridium difficile* Infection, Finland, 1996-2004. *Emerging Infectious Diseases*, 15, 761-765.
- MACCALLUM, R. M., MARTIN, A. C. R. & THORNTON, J. M. (1996) Antibody-antigen Interactions: Contact Analysis and Binding Site Topography. *Journal of Molecular Biology*, 262, 732-745.
- MAENG, B., NAM, D. & KIM, Y. Coexpression of molecular chaperones to enhance functional expression of anti-BNP scFv in the cytoplasm of *Escherichia coli*; for the detection of B-type natriuretic peptide. *World Journal of Microbiology and Biotechnology*, 27, 1391-1398.
- MAKOWSKI, L. (1993) Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene*, 128, 5-11.
- MALECKI, M., HSU, A., TRUONG, L. & SANCHEZ, S. (2002) Molecular immunolabeling with recombinant single-chain variable fragment (scFv) antibodies designed with metal-binding domains. *Proceedings of the National Academy of Sciences*, 99, 213-218.
- MALEMBIC, S., SAILLARD, C., BOVE, J. M. & GARNIER, M. (2002) Effect of Polyclonal, Monoclonal, and Recombinant (Single-Chain Variable Fragment) Antibodies on In Vitro Morphology, Growth, and Metabolism of the Phytopathogenic Mollicute *Spiroplasma citri*. *Appl. Environ. Microbiol.*, 68, 2113-2119.
- MALMBORG, A.-C., DUEÑAS, M., OHLIN, M., SDERLIND, E. & BORREBAECK, C. A. K. (1996) Selection of binders from phage displayed antibody libraries using the BIAcore(TM) biosensor. *Journal of Immunological Methods*, 198, 51-57.
- MANUELA BIANCO, G. F., ADRIANO QUATTRINI, PATRIZIA SPIGAGLIA, FABRIZIO BARBANTI, PAOLA MASTRANTONIO, CLARA M. AUSIELLO* (2011) **Immunomodulatory activities of surface layer proteins (SLPs) 1 obtained from epidemic and hypervirulent *Clostridium difficile* strains.** *Journal of Medical Microbiology*,
- MARKS, J. D., HOOGENBOOM, H. R., BONNERT, T. P., MCCAFFERTY, J., GRIFFITHS, A. D. & WINTER, G. (1991) By-passing immunization : Human antibodies from V-gene libraries displayed on phage. *Journal of Molecular Biology*, 222, 581-597.
- MASUDA, K., ITOH, M. & KAWATA, T. (1989) CHARACTERIZATION AND REASSEMBLY OF A REGULAR ARRAY IN THE CELL-WALL OF CLOSTRIDIUM-DIFFICILE GAI-4131. *Microbiology and Immunology*, 33, 287-298.
- MATTHEWS, D. & WELLS, J. (1993) Substrate phage: selection of protease substrates by monovalent phage display. *Science*, 260, 1113-1117.
- MAURI, P. L., PIETTA, P. G., MAGGIONI, A., CERQUETTI, M., SEBASTIANELLI, A. & MASTRANTONIO, P. (1999) Characterization of surface layer proteins from *Clostridium difficile* by liquid chromatography/electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom*, 13, 695-703.
- MAZMANIAN, S. K., LIU, G., TON-THAT, H. & SCHNEEWIND, O. (1999) Staphylococcus aureus sortase, an enzyme that anchors surface proteins to the cell wall. *Science*, 285, 760-3.
- MAZMANIAN, S. K., TON-THAT, H., SU, K. & SCHNEEWIND, O. (2002) An iron-regulated sortase anchors a class of surface protein during Staphylococcus aureus pathogenesis. *Proc Natl Acad Sci U S A*, 99, 2293-8.
- MCCAFFERTY, J. & GLOVER, D. R. (2000) Engineering therapeutic proteins. *Current Opinion in Structural Biology*, 10, 417-420.
- MCCAFFERTY, J., GRIFFITHS, A. D., WINTER, G. & CHISWELL, D. J. (1990a) PHAGE ANTIBODIES - FILAMENTOUS PHAGE DISPLAYING ANTIBODY VARIABLE DOMAINS. *Nature*, 348, 552-554.
- MCCAFFERTY, J., GRIFFITHS, A. D., WINTER, G. & CHISWELL, D. J. (1990b) Phage antibodies: filamentous phage displaying antibody variable domains. *Nature*, 348, 552-554.
- MCCONNELL, S. J., KENDALL, M. L., REILLY, T. M. & HOESS, R. H. (1994) Constrained peptide libraries as a tool for finding mimotopes. *Gene*, 151, 115-118.
- MCCOUBREY, J. & POXTON, I. R. (2001) Variation in the surface layer proteins of *Clostridium difficile*. *FEMS Immunol Med Microbiol*, 31, 131-5.
- MCDONNELL, P. A., SHON, K., KIM, Y. & OPELLA, S. J. (1993) fd Coat Protein Structure in Membrane Environments. *Journal of Molecular Biology*, 233, 447-463.
- MCFARLAND, L. V., BENEDA, H. W., CLARRIDGE, J. E. & RAUGI, G. J. (2007) Implications of the changing face of *Clostridium difficile* disease for health care practitioners. *Am J Infect Control*, 35, 237-53.

- MCFARLAND, L. V., ELMER, G. W. & SURAWICZ, C. M. (2002) Breaking the cycle: Treatment strategies for 163 cases of recurrent *Clostridium difficile* disease. *American Journal of Gastroenterology*, *97*, 1769-1775.
- MCGUINNESS, B. T., WALTER, G., FITZGERALD, K., SCHULER, P., MAHONEY, W., DUNCAN, A. R. & HOOGENBOOM, H. R. (1996) Phage diabody repertoires for selection of large numbers of bispecific antibody fragments. *Nat Biotech*, *14*, 1149-1154.
- MECHALY, A., ZAHAVY, E. & FISHER, M. (2008) Development and Implementation of a Single-Chain Fv Antibody for Specific Detection of *Bacillus anthracis* Spores. *Appl. Environ. Microbiol.*, *74*, 818-822.
- MELKANI, G. C., SIELAFF, R., ZARDENETA, G. & MENDOZA, J. A. Interaction of oxidized chaperonin GroEL with an unfolded protein at low temperatures. *Bioscience reports*, *32*, 299-303.
- MINTZ, P. J., KIM, J., DO, K.-A., WANG, X., ZINNER, R. G., CRISTOFANILLI, M., ARAP, M. A., HONG, W. K., TRONCOSO, P., LOGOTHETIS, C. J., PASQUALINI, R. & ARAP, W. (2003) Fingerprinting the circulating repertoire of antibodies from cancer patients. *Nat Biotech*, *21*, 57-63.
- MODENA, S., GOLLAMUDI, S. & FRIEDENBERG, F. (2006) Continuation of antibiotics is associated with failure of metronidazole for *Clostridium difficile*-associated diarrhea. *Journal of Clinical Gastroenterology*, *40*, 49-54.
- MUSHER, D. M., ASLAM, S., LOGAN, N., NALLACHERU, S., BHAILA, I., BORCHERT, F. & HAMILL, R. J. (2005) Relatively poor outcome after treatment of *Clostridium difficile* colitis with metronidazole. *Clinical Infectious Diseases*, *40*, 1586-1590.
- MUSSER, J. M., HAUSER, A. R., KIM, M. H., SCHLIEVERT, P. M., NELSON, K. & SELANDER, R. K. (1991) *Streptococcus pyogenes* causing toxic-shock-like syndrome and other invasive diseases: clonal diversity and pyrogenic exotoxin expression. *Proc Natl Acad Sci U S A*, *88*, 2668-72.
- MUTO, C. A., POKRYWKA, M., SHUTT, K., MENDELSON, A. B., NOURI, K., POSEY, K., ROBERTS, T., CROYLE, K., KRYSOFIAK, S., PATEL-BROWN, S., PASCULLE, A. W., PATERSON, D. L. & SAUL, M. (2005) A large outbreak of *Clostridium Difficile* - Associated disease with an unexpected proportion of deaths and colectomies at a teaching hospital following increased fluoroquinolone use. *Infection Control and Hospital Epidemiology*, *26*, 273-280.
- MYLONAKIS, E., RYAN, E. T. & CALDERWOOD, S. B. (2001) *Clostridium difficile*-associated diarrhea - A review. *Archives of Internal Medicine*, *161*, 525-533.
- NACHNANI, J. S., BULCHANDANI, D. & ALLEN, M. J. (2008) Proton pump inhibitors are an independent risk factor for an increased length of hospital stay in patients with *Clostridium difficile* infection. *Indian J Gastroenterol*, *27*, 171-2.
- NAKAMURA, S., MIKAWA, M., NAKASHIO, S., TAKABATAKE, M., OKADO, I., YAMAKAWA, K., SERIKAWA, T., OKUMURA, S. & NISHIDA, S. (1981) ISOLATION OF CLOSTRIDIUM-DIFFICILE FROM THE FECES AND THE ANTIBODY IN SERA OF YOUNG AND ELDERLY ADULTS. *Microbiology and Immunology*, *25*, 345-351.
- NEUBERGER, M. (1996) Generating high-avidity human Mabs in mice. *Nat Biotech*, *14*, 826-826.
- O'BRIEN, J. B., MCCABE, M. S., ATHIE-MORALES, V., MCDONALD, G. S., NI EIDHIN, D. B. & KELLEHER, D. P. (2005) Passive immunisation of hamsters against *Clostridium difficile* infection using antibodies to surface layer proteins. *FEMS Microbiol Lett*, *246*, 199-205.
- OHKAWA, I. & WEBSTER, R. E. (1981) The orientation of the major coat protein of bacteriophage f1 in the cytoplasmic membrane of *Escherichia coli*. *Journal of Biological Chemistry*, *256*, 9951-9958.
- OHKURA, T., KIKUCHI, Y., KONO, N., ITAMURA, S., KOMASE, K., MOMOSE, F. & MORIKAWA, Y. Epitope mapping of neutralizing monoclonal antibody in avian influenza A H5N1 virus hemagglutinin. *Biochemical and Biophysical Research Communications*, *418*, 38-43.
- OLDENBURG, K. R., LOGANATHAN, D., GOLDSTEIN, I. J., SCHULTZ, P. G. & GALLOP, M. A. (1992) Peptide ligands for a sugar-binding protein isolated from a random peptide library. *Proceedings of the National Academy of Sciences*, *89*, 5393-5397.
- ONEILL, G., ADAMS, J. E., BOWMAN, R. A. & RILEY, T. V. (1993) A MOLECULAR CHARACTERIZATION OF CLOSTRIDIUM-DIFFICILE ISOLATES FROM HUMANS, ANIMALS AND THEIR ENVIRONMENTS. *Epidemiology and Infection*, *111*, 257-264.
- ONEILL, G. L., BEAMAN, M. H. & RILEY, T. V. (1991) RELAPSE VERSUS REINFECTION WITH CLOSTRIDIUM-DIFFICILE. *Epidemiology and Infection*, *107*, 627-635.
- PANTOSTI, A., CERQUETTI, M., VITI, F., ORTISI, G. & MASTRANTONIO, P. (1989a) IMMUNOBLOT ANALYSIS OF SERUM IMMUNOGLOBULIN-G RESPONSE TO SURFACE-PROTEINS OF CLOSTRIDIUM-DIFFICILE IN PATIENTS WITH ANTIBIOTIC-ASSOCIATED DIARRHEA. *Journal of Clinical Microbiology*, *27*, 2594-2597.
- PANTOSTI, A., CERQUETTI, M., VITI, F., ORTISI, G. & MASTRANTONIO, P. (1989b) Immunoblot analysis of serum immunoglobulin G response to surface proteins of *Clostridium difficile* in patients with antibiotic-associated diarrhea. *J. Clin. Microbiol.*, *27*, 2594-2597.

- PARREN, P. W. H. I. & BURTON, D. R. (2001) The antiviral activity of antibodies in vitro and in vivo. *Advances in Immunology*. Academic Press.
- PASQUALINI, R. & RUOSLAHTI, E. (1996) Organ targeting In vivo using phage display peptide libraries. *Nature*, 380, 364-366.
- PATERSON, G. K. & MITCHELL, T. J. (2004) The biology of Gram-positive sortase enzymes. *Trends in Microbiology*, 12, 89-95.
- PECHINE, S., GLEIZES, A., JANOIR, C., GORGES-KERGOT, R., BARC, M. C., DELMEE, M. & COLLIGNON, A. (2005a) Immunological properties of surface proteins of *Clostridium difficile*. *J Med Microbiol*, 54, 193-6.
- PECHINE, S., JANOIR, C., BOUREAU, H., GLEIZES, A., TSAPIS, N., HOYS, S., FATTAL, E. & COLLIGNON, A. (2007) Diminished intestinal colonization by *Clostridium difficile* and immune response in mice after mucosal immunization with surface proteins of *Clostridium difficile*. *Vaccine*, 25, 3946-54.
- PECHINE, S., JANOIR, C. & COLLIGNON, A. (2005b) Variability of *Clostridium difficile* surface proteins and specific serum antibody response in patients with *Clostridium difficile*-associated disease. *J Clin Microbiol*, 43, 5018-25.
- PEPIN, J., VALIQUETTE, L., ALARY, M. E., VILLEMURE, P., PELLETIER, A., FORGET, K., PEPIN, K. & CHOUINARD, D. (2004) *Clostridium difficile*-associated diarrhea in a region of Quebec from 1991 to 2003: a changing pattern of disease severity. *Canadian Medical Association Journal*, 171, 466-472.
- PETERSON, L. R., MANSON, R. U., PAULE, S. M., HACEK, D. M., ROBICSEK, A., THOMSON, R. B., JR. & KAUL, K. L. (2007) Detection of toxigenic *Clostridium difficile* in stool samples by real-time polymerase chain reaction for the diagnosis of *C. difficile*-associated diarrhea. *Clin Infect Dis*, 45, 1152-60.
- PITUCH, H., BAKKER, D., KUIJPER, E., OBUCH-WOSZCZATYNSKI, P., WULTANSKA, D., NURZYNSKA, G., BIELEC, A., BAR-ANDZIAK, E. & LUCZAK, M. (2008) First isolation of *Clostridium difficile* PCR-ribotype 027/toxinotype III in Poland. *Polish Journal of Microbiology*, 57, 267-268.
- PLÜCKTHUN, A., SKERRA, A. & JOHN, J. L. (1989) [34] Expression of functional antibody Fv and Fab fragments in *Escherichia coli*. *Methods in Enzymology*. Academic Press.
- PRINCE, A. (2006) Flagellar Activation of Epithelial Signaling. *American Journal of Respiratory Cell and Molecular Biology*, 34, 548-551.
- QAZI, O., HITCHEN, P., TISSOT, B., PANICO, M., MORRIS, H. R., DELL, A. & FAIRWEATHER, N. (2009) Mass spectrometric analysis of the S-layer proteins from *Clostridium difficile* demonstrates the absence of glycosylation. *Journal of Mass Spectrometry*, 44, 368-374.
- RAKONJAC, J., FENG, J. N. & MODEL, P. (1999) Filamentous phage are released from the bacterial membrane by a two-step mechanism involving a short C-terminal fragment of pIII. *Journal of Molecular Biology*, 289, 1253-1265.
- RAPOZA, M. P. & WEBSTER, R. E. (1993) The filamentous bacteriophage assembly proteins require the bacterial SecA protein for correct localization to the membrane. *J. Bacteriol.*, 175, 1856-1859.
- REICHERT, J. M. (2001) Monoclonal antibodies in the clinic. *Nat Biotech*, 19, 819-822.
- REICHERT, J. M. & VALGE-ARCHER, V. E. (2007) Development trends for monoclonal antibody cancer therapeutics. *Nat Rev Drug Discov*, 6, 349-356.
- REISNER, Y. & DAGAN, S. (1998) The Trimer mouse: generating human monoclonal antibodies and an animal model for human diseases. *Trends in Biotechnology*, 16, 242-246.
- RELLER, M. E., LEMA, C. A., PERL, T. M., CAI, M., ROSS, T. L., SPECK, K. A. & CARROLL, K. C. (2007) Yield of Stool Culture with Isolate Toxin Testing versus a Two-Step Algorithm Including Stool Toxin Testing for the Detection of Toxigenic *Clostridium difficile*. *J Clin Microbiol*.
- REN, Z. J., TIAN, C. J., ZHU, Q. S., ZHAO, M. Y., XIN, A. G., NIE, W. X., LING, S. R., ZHU, M. W., WU, J. Y., LAN, H. Y., CAO, Y. C. & BI, Y. Z. (2008) Orally delivered foot-and-mouth disease virus capsid protomer vaccine displayed on T4 bacteriophage surface: 100% protection from potency challenge in mice. *Vaccine*, 26, 1471-1481.
- RHEE, K. J., JASPER, P. J., SETHUPATHI, P., SHANMUGAM, M., LANNING, D. & KNIGHT, K. L. (2005) Positive selection of the peripheral B cell repertoire in gut-associated lymphoid tissues. *J Exp Med*, 201, 55-62.
- RHEE, K. J., SETHUPATHI, P., DRIKS, A., LANNING, D. K. & KNIGHT, K. L. (2004) Role of commensal bacteria in development of gut-associated lymphoid tissues and preimmune antibody repertoire. *J Immunol*, 172, 1118-24.
- RHYNER, C., WEICHEL, M., FLÜCKIGER, S., HEMMANN, S., KLEBER-JANKE, T. & CRAMERI, R. (2004) Cloning allergens via phage display. *Methods*, 32, 212-218.
- RIECHMANN, L., CLARK, M., WALDMANN, H. & WINTER, G. (1988) Reshaping human antibodies for therapy. *Nature*, 332, 323-327.
- RIEGLER, M., SEDIVY, R., POTHOUKAKIS, C., HAMILTON, G., ZACHERI, J., BISCHOF, G., COSENTINI, E., FEIL, W., SCHIESSEL, R., LAMONT, J. T. & WENZL, E. (1995) CLOSTRIDIUM-DIFFICILE TOXIN-B IS MORE

- POTENT THAN TOXIN-A IN DAMAGING HUMAN COLONIC EPITHELIUM IN-VITRO. *Journal of Clinical Investigation*, 95, 2004-2011.
- ROBERTS, B. L., MARKLAND, W., SIRANOSIAN, K., SAXENA, M. J., GUTERMAN, S. K. & LADNER, R. C. (1992) Protease inhibitor display M13 phage: selection of high-affinity neutrophil elastase inhibitors. *Gene*, 121, 9-15.
- RODRIGUEZ-PALACIOS, A., STAMPFLI, H. R., DUFFIELD, T., PEREGRINE, A. S., TROTZ-WILLIAMS, L. A., ARROYO, L. G., BRAZIER, J. S. & WEESE, J. S. (2006) Clostridium difficile PCR ribotypes in calves, Canada. *Emerging Infectious Diseases*, 12, 1730-1736.
- RODRIGUEZ, P. L. & CARRASCO, L. (1995) Improved factor Xa cleavage of fusion proteins containing maltose binding protein. *Biotechniques*, 18, 238, 241-3.
- ROTIMI, V. O., JAMAL, W. Y., MOKADDAS, E. M., BRAZIER, J. S., JOHNY, M. & DUERDEN, B. I. (2003) Prevalent PCR ribotypes of clinical and environmental strains of Clostridium difficile isolated from intensive-therapy unit patients in Kuwait. *Journal of Medical Microbiology*, 52, 705-709.
- RUBERTI, F., CATTANEO, A. & BRADBURY, A. (1994) The use of the RACE method to clone hybridoma cDNA when V region primers fail. *Journal of Immunological Methods*, 173, 33-39.
- RUBIN, M. S., BODENSTEIN, L. E. & KENT, K. C. (1995) SEVERE CLOSTRIDIUM-DIFFICILE COLITIS. *Diseases of the Colon & Rectum*, 38, 350-354.
- RUSSEL, M., WHIRLOW, H., SUN, T. P. & WEBSTER, R. E. (1988) LOW-FREQUENCY INFECTION OF F-BACTERIA BY TRANSDUCING PARTICLES OF FILAMENTOUS BACTERIOPHAGES. *Journal of Bacteriology*, 170, 5312-5316.
- SANCHEZ-HURTADO, K. & POXTON, I. R. (2008) Enhancement of the cytotoxic activity of Clostridium difficile toxin A by surface-associated antigens. *J Med Microbiol*, 57, 739-44.
- SANCHEZ, T. H., BROOKS, J. T., SULLIVAN, P. S., JUHASZ, M., MINTZ, E., DWORKIN, M. S., JONES, J. L. & ADULT ADOLESCENT SPECTRUM, H. I. V. D. (2005) Bacterial diarrhea in persons with HIV infection, United States, 1992-2002. *Clinical Infectious Diseases*, 41, 1621-1627.
- SAVARIAU-LACOMME, M.-P., LEBARBIER, C., KARJALAINEN, T., COLLIGNON, A. & JANOIR, C. (2003) Transcription and Analysis of Polymorphism in a Cluster of Genes Encoding Surface-Associated Proteins of Clostridium difficile. *J. Bacteriol.*, 185, 4461-4470.
- SAWABE, E., KATO, H., OSAWA, K., CHIDA, T., TOJO, N., ARAKAWA, Y. & OKAMURA, N. (2007) Molecular analysis of Clostridium difficile at a university teaching hospital in Japan: a shift in the predominant type over a five-year period. *European Journal of Clinical Microbiology & Infectious Diseases*, 26, 695-703.
- SCOTT, J. K. & SMITH, G. P. (1990) SEARCHING FOR PEPTIDE LIGANDS WITH AN EPITOPE LIBRARY. *Science*, 249, 386-390.
- SCOTT, J. R. & BARNETT, T. C. (2006) Surface proteins of gram-positive bacteria and how they get there. *Annu Rev Microbiol*, 60, 397-423.
- SCOTT, N., QAZI, O., WRIGHT, M. J., FAIRWEATHER, N. F. & DEONARAIN, M. P. Characterisation of a panel of anti-tetanus toxin single-chain Fvs reveals cooperative binding. *Mol Immunol*, 47, 1931-41.
- SEBAIHIA, M., WREN, B. W., MULLANY, P., FAIRWEATHER, N. F., MINTON, N., STABLER, R., THOMSON, N. R., ROBERTS, A. P., CERDENO-TARRAGA, A. M., WANG, H., HOLDEN, M. T., WRIGHT, A., CHURCHER, C., QUAIL, M. A., BAKER, S., BASON, N., BROOKS, K., CHILLINGWORTH, T., CRONIN, A., DAVIS, P., DOWD, L., FRASER, A., FELTWELL, T., HANCE, Z., HOLROYD, S., JAGELS, K., MOULE, S., MUNGALL, K., PRICE, C., RABBINOWITSCH, E., SHARP, S., SIMMONDS, M., STEVENS, K., UNWIN, L., WHITHEAD, S., DUPUY, B., DOUGAN, G., BARRELL, B. & PARKHILL, J. (2006) The multidrug-resistant human pathogen Clostridium difficile has a highly mobile, mosaic genome. *Nat Genet*, 38, 779-86.
- SEVERSON, K. M., MALLOZZI, M., DRIKS, A. & KNIGHT, K. L. B cell development in GALT: role of bacterial superantigen-like molecules. *J Immunol*, 184, 6782-9.
- SHARP, J. & POXTON, I. R. (1988) THE CELL-WALL PROTEINS OF CLOSTRIDIUM-DIFFICILE. *Fems Microbiology Letters*, 55, 99-103.
- SHEARMAN, C., POLLOCK, D., WHITE, G., HEHIR, K., MOORE, G., KANZY, E. & KURRLE, R. (1991) Construction, expression and characterization of humanized antibodies directed against the human alpha/beta T cell receptor. *The Journal of Immunology*, 147, 4366-4373.
- SHENGFENG, C., PING, L., TAO, S., XIN, W. & GUOFENG, W. (2003) Construction, Expression, Purification, Refold and Activity Assay of a Specific ScFv Fragment against Foot and Mouth Disease Virus. *Veterinary Research Communications*, 27, 243-256.
- SHIN, B. M., KUAK, E. Y., YOO, H. M., KIM, E. C., LEE, K., KANG, J. O., WHANG, D. H. & SHIN, J. H. (2008) Multicentre study of the prevalence of toxigenic Clostridium difficile in Korea: results of a retrospective study 2000-2005. *Journal of Medical Microbiology*, 57, 697-701.

- SIMANGO, C. & MWAKURUDZA, S. (2008) Clostridium difficile in broiler chickens sold at market places in Zimbabwe and their antimicrobial susceptibility. *International Journal of Food Microbiology*, 124, 268-270.
- SINGH, P. K., AGRAWAL, R., KAMBOJ, D. V., GUPTA, G., BOOPATHI, M., GOEL, A. K. & SINGH, L. Construction of a Single-Chain Variable-Fragment Antibody against the Superantigen Staphylococcal Enterotoxin B. *Appl. Environ. Microbiol.*, 76, 8184-8191.
- SKERRA, A. (1993) BACTERIAL EXPRESSION OF IMMUNOGLOBULIN FRAGMENTS. *Current Opinion in Immunology*, 5, 256-262.
- SKERRA, A. & PLUCKTHUN, A. (1988a) ASSEMBLY OF A FUNCTIONAL IMMUNOGLOBULIN-FV FRAGMENT IN ESCHERICHIA-COLI. *Science*, 240, 1038-1041.
- SKERRA, A. & PLUCKTHUN, A. (1988b) Assembly of a functional immunoglobulin Fv fragment in Escherichia coli. *Science*, 240, 1038-1041.
- SMITH, A. (2005) Outbreak of Clostridium difficile infection in an English hospital linked to hypervirulent strains in Canada and the US. *Euro Surveill*, 10, E050630.2.
- SMITH, G. (1985a) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, 228, 1315-1317.
- SMITH, G. P. (1985b) FILAMENTOUS FUSION PHAGE - NOVEL EXPRESSION VECTORS THAT DISPLAY CLONED ANTIGENS ON THE VIRION SURFACE. *Science*, 228, 1315-1317.
- SMITH, G. P. & PETRENKO, V. A. (1997) Phage display. *Chemical Reviews*, 97, 391-410.
- SOMERVILLE, J. E., GOSHORN, S. C., FELL, H. P. & DARVEAU, R. P. (1994) BACTERIAL ASPECTS ASSOCIATED WITH THE EXPRESSION OF A SINGLE-CHAIN ANTIBODY FRAGMENT IN ESCHERICHIA-COLI. *Applied Microbiology and Biotechnology*, 42, 595-603.
- SONGER, J. G., TRINH, H. T., KILLGORE, G. E., THOMPSON, A. D., MCDONALD, L. C. & LIMBAGO, B. M. (2009) Clostridium difficile in Retail Meat Products, USA, 2007. *Emerging Infectious Diseases*, 15, 819-821.
- SOUGIOLTZIS, S., KYNE, L., DRUDY, D., KEATES, S., MAROO, S., POTHOUKAKIS, C., GIANNASCA, P., WARNY, M., LEE, C., MONATH, T. & KELLY, C. P. (2004) Clostridium difficile toxoid vaccine in recurrent C. difficile associated diarrhea. *Gastroenterology*, 126, A512-A512.
- SPIGAGLIA, P., BARBANTI, F. & MASTRANTONIO, P. Surface Layer Protein A Variant of Clostridium difficile PCR-Ribotype 027. *Emerging Infectious Diseases*, 17, 317-319.
- SPIGAGLIA, P., GALEOTTI, C. L., BARBANTI, F., SCARSELLI, M., VAN BROECK, J. & MASTRANTONIO, P. The LMW surface-layer proteins of Clostridium difficile PCR ribotypes 027 and 001 share common immunogenic properties. *Journal of Medical Microbiology*, 60, 1168-1173.
- STABLER, R. A., HE, M., DAWSON, L., MARTIN, M., VALIENTE, E., CORTON, C., LAWLEY, T. D., SEBAIHIA, M., QUAIL, M. A., ROSE, G., GERDING, D. N., GIBERT, M., POPOFF, M. R., PARKHILL, J., DOUGAN, G. & WREN, B. W. (2009) Comparative genome and phenotypic analysis of Clostridium difficile 027 strains provides insight into the evolution of a hypervirulent bacterium. *Genome Biology*, 10, 15.
- STEEN, A., BUIST, G., LEENHOUTS, K. J., EL KHATTABI, M., GRIJPSTRA, F., ZOMER, A. L., VENEMA, G., KUIPERS, O. P. & KOK, J. (2003) Cell wall attachment of a widely distributed peptidoglycan binding domain is hindered by cell wall constituents. *J Biol Chem*, 278, 23874-81.
- SUNDRAM, F., GUYOT, A., CARBOO, I., GREEN, S., LILAONITKUL, M. & SCOURFIELD, A. (2009) Clostridium difficile ribotypes 027 and 106: clinical outcomes and risk factors. *Journal of Hospital Infection*, 72, 111-118.
- SURAWICZ, C. M., MCFARLAND, L. V., GREENBERG, R. N., RUBIN, M., FEKETY, R., MULLIGAN, M. E., GARCIA, R. J., BRANDMARKER, S., BOWEN, K., BORJAL, D. & ELMER, G. W. (2000) The search for a better treatment for recurrent Clostridium difficile disease: Use of high-dose vancomycin combined with Saccharomyces boulardii. *Clinical Infectious Diseases*, 31, 1012-1017.
- SWINDELLS, J., BRENWALD, N., READING, N. & OPPENHEIM, B. Evaluation of Diagnostic Tests for Clostridium difficile Infection. *Journal of Clinical Microbiology*, 48, 606-608.
- TAE, C. H., JUNG, S. A., SONG, H. J., KIM, S. E., CHOI, H. J., LEE, M., HWANG, Y., KIM, H. & LEE, K. (2009) The First Case of Antibiotic-associated Colitis by Clostridium difficile PCR Ribotype 027 in Korea. *Journal of Korean Medical Science*, 24, 520-524.
- TAKEOKA, A., TAKUMI, K., KOGA, T. & KAWATA, T. (1991) PURIFICATION AND CHARACTERIZATION OF S-LAYER PROTEINS FROM CLOSTRIDIUM-DIFFICILE GAI-0714. *Journal of General Microbiology*, 137, 261-267.
- TASTREYRE, A., BARC, M. C., COLLIGNON, A., BOUREAU, H. & KARJALAINEN, T. (2001a) Role of FliC and FliD flagellar proteins of Clostridium difficile in adherence and gut colonization. *Infect Immun*, 69, 7937-40.
- TASTREYRE, A., BARC, M. C., KARJALAINEN, T., DODSON, P., HYDE, S., BOURLILOUX, P. & BORRIELLO, P. (2000a) A Clostridium difficile gene encoding flagellin. *Microbiology*, 146 (Pt 4), 957-66.

- TASTEYRE, A., KARJALAINEN, T., AVESANI, V., DELMEE, M., COLLIGNON, A., BOURLIOUX, P. & BARC, M. C. (2000b) Phenotypic and genotypic diversity of the flagellin gene (fliC) among *Clostridium difficile* isolates from different serogroups. *J Clin Microbiol*, 38, 3179-86.
- TASTEYRE, A., KARJALAINEN, T., AVESANI, V., DELMEE, M., COLLIGNON, A., BOURLIOUX, P. & BARC, M. C. (2001b) Molecular characterization of fliD gene encoding flagellar cap and its expression among *Clostridium difficile* isolates from different serogroups. *J Clin Microbiol*, 39, 1178-83.
- TEDESCO, F. J., BARTON, R. W. & ALPERS, D. H. (1974) CLINDAMYCIN-ASSOCIATED COLITIS - PROSPECTIVE STUDY. *Annals of Internal Medicine*, 81, 429-433.
- TERHES, G., URBAN, E., KONKOLY-THEGE, M., SZEKELY, E., BRAZIER, J. S., KUIJPER, E. J. & NAGY, E. (2009) First isolation of *Clostridium difficile* PCR ribotype 027 from a patient with severe persistent diarrhoea in Hungary. *Clinical Microbiology and Infection*, 15, 885-886.
- TODD, M. J., VIITANEN, P. V. & LORIMER, G. H. (1993) HYDROLYSIS OF ADENOSINE 5'-TRIPHOSPHATE BY ESCHERICHIA-COLI GROEL - EFFECTS OF GROES AND POTASSIUM-ION. *Biochemistry*, 32, 8560-8567.
- TOMA, S., LESIAK, G., MAGUS, M., LO, H. L. & DELMEE, M. (1988) Serotyping of *Clostridium difficile*. *J Clin Microbiol*, 26, 426-8.
- TOMIZUKA, K., SHINOHARA, T., YOSHIDA, H., UEJIMA, H., OHGUMA, A., TANAKA, S., SATO, K., OSHIMURA, M. & ISHIDA, I. (2000) Double trans-chromosomal mice: Maintenance of two individual human chromosome fragments containing Ig heavy and λ loci and expression of fully human antibodies. *Proceedings of the National Academy of Sciences*, 97, 722-727.
- TRAMONTANO, A., CHOTHIA, C. & LESK, A. M. (1990) Framework residue 71 is a major determinant of the position and conformation of the second hypervariable region in the VH domains of immunoglobulins. *Journal of Molecular Biology*, 215, 175-182.
- TWINE, S. M., REID, C. W., AUBRY, A., MCMULLIN, D. R., FULTON, K. M., AUSTIN, J. & LOGAN, S. M. (2009) Motility and Flagellar Glycosylation in *Clostridium difficile*. *Journal of Bacteriology*, 191.
- VAN BOCKSTAELE, F., HOLZ, J. B. & REVETS, H. (2009) The development of nanobodies for therapeutic applications. *Curr Opin Investig Drugs*, 10, 1212-24.
- VAN DEN BERG, R. J., SCHAAP, I., TEMPLETON, K. E., KLAASSEN, C. H. W. & KUIJPER, E. J. (2007) Typing and Subtyping of *Clostridium difficile* Isolates by Using Multiple-Locus Variable-Number Tandem-Repeat Analysis. *Journal of Clinical Microbiology*, 45, 1024-1028.
- VAN DISSEL, J. T., DE GROOT, N., HENSGENS, C. M. H., NUMAN, S., KUIJPER, E. J., VELDKAMP, P. & VAN'T WOUT, J. (2005) Bovine antibody-enriched whey to aid in the prevention of a relapse of *Clostridium difficile* associated diarrhoea: preclinical and preliminary clinical data. *Journal of Medical Microbiology*, 54, 197-205.
- VAN HOUTEN, N. E., ZWICK, M. B., MENENDEZ, A. & SCOTT, J. K. (2006) Filamentous phage as an immunogenic carrier to elicit focused antibody responses against a synthetic peptide. *Vaccine*, 24, 4188-4200.
- VANEWIJK, W., DEKRUIF, J., GERMERAAD, W. T. V., BERENDES, P., ROPKE, C., PLATENBURG, P. P. & LOGTENBERG, T. (1997) Subtractive isolation of phage-displayed single-chain antibodies to thymic stromal cells by using intact thymic fragments. *Proceedings of the National Academy of Sciences of the United States of America*, 94, 3903-3908.
- VANWEZENBEEK, P., HULSEBOS, T. J. M. & SCHOENMAKERS, J. G. G. (1980) NUCLEOTIDE-SEQUENCE OF THE FILAMENTOUS BACTERIOPHAGE M13-DNA GENOME - COMPARISON WITH PHAGE-FD. *Gene*, 11, 129-148.
- VAUGHAN, T. J., WILLIAMS, A. J., PRITCHARD, K., OSBOURN, J. K., POPE, A. R., EARNSHAW, J. C., MCCAFFERTY, J., HODITS, R. A., WILTON, J. & JOHNSON, K. S. (1996) Human Antibodies with Sub-nanomolar Affinities Isolated from a Large Non-immunized Phage Display Library. *Nat Biotech*, 14, 309-314.
- VIEIRA, J., MESSING, J. & RAY WU, L. G. (1987) [1] Production of single-stranded plasmid DNA. *Methods in Enzymology*. Academic Press.
- VIRNEKAS, B., GE, L., PLUÛ, -KTHUN, A., SCHNEIDER, K. C., WELLNHOFER, G. N. & MORONEY, S. E. (1994) Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Research*, 22, 5600-5607.
- VOHRA, P. & POXTON, I. R. Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. *Microbiology*, 157, 1343-1353.
- VOHRA, P. & POXTON, I. R. Comparison of toxin and spore production in clinically relevant strains of *Clostridium difficile*. *Microbiology-Sgm*, 157, 1343-1353.
- VOLLMER, W., JORIS, B., CHARLIER, P. & FOSTER, S. (2008) Bacterial peptidoglycan (murein) hydrolases. *Fems Microbiology Reviews*, 32, 259-286.

- VOTH, D. E. & BALLARD, J. D. (2005) Clostridium difficile toxins: mechanism of action and role in disease. *Clin Microbiol Rev*, 18, 247-63.
- WALIGORA, A. J., BARC, M. C., BOURLIOUX, P., COLLIGNON, A. & KARJALAINEN, T. (1999) Clostridium difficile cell attachment is modified by environmental factors. *Appl Environ Microbiol*, 65, 4234-8.
- WALIGORA, A. J., HENNEQUIN, C., MULLANY, P., BOURLIOUX, P., COLLIGNON, A. & KARJALAINEN, T. (2001) Characterization of a cell surface protein of Clostridium difficile with adhesive properties. *Infect Immun*, 69, 2144-53.
- WANDERSMAN, C. (1989) Secretion, processing and activation of bacterial extracellular proteases. *Mol Microbiol*, 3, 1825-31.
- WARD, R. L., CLARK, M. A., LEES, J. & HAWKINS, N. J. (1996) Retrieval of human antibodies from phage-display libraries using enzymatic cleavage. *Journal of Immunological Methods*, 189, 73-82.
- WARNY, M., PEPIN, J., FANG, A., KILLGORE, G., THOMPSON, A., BRAZIER, J., FROST, E. & MCDONALD, L. C. (2005a) Toxin production by an emerging strain of Clostridium difficile associated with outbreaks of severe disease in North America and Europe. *Lancet*, 366, 1079-1084.
- WARNY, M., PEPIN, J., FANG, A., KILLGORE, G., THOMPSON, A., BRAZIER, J., FROST, E. & MCDONALD, L. C. (2005b) Toxin production by an emerging strain of Clostridium difficile associated with outbreaks of severe disease in North America and Europe. *Lancet*, 366, 1079-84.
- WATKINS, N. A. & OUWEHAND, W. H. (2000) Introduction to Antibody Engineering and Phage Display. *Vox Sanguinis*, 78, 72-79.
- WEESE, J. S., AVERY, B. P., ROUSSEAU, J. & REID-SMITH, R. J. (2009) Detection and enumeration of Clostridium difficile spores in retail beef and pork. *Applied and Environmental Microbiology*, 75, 5009-5011.
- WELSCHOF, M., TERNESS, P., KIPRIYANOV, S. M., STANESCU, D., BREITLING, F., DÍRSAM, H., DÍBEL, S., LITTLE, M. & OPELZ, G. (1997) The antigen-binding domain of a human IgG-anti-F(ab²)₂ autoantibody. *Proceedings of the National Academy of Sciences*, 94, 1902-1907.
- WESOLOWSKI, J., ALZOGARAY, V., REYELT, J., UNGER, M., JUAREZ, K., URRUTIA, M., CAUERHFF, A., DANQUAH, W., RISSIEK, B., SCHEUPLEIN, F., SCHWARZ, N., ADRIOUCH, S., BOYER, O., SEMAN, M., LICEA, A., SERREZE, D. V., GOLDBAUM, F. A., HAAG, F. & KOCH-NOLTE, F. (2009) Single domain antibodies: promising experimental and therapeutic tools in infection and immunity. *Med Microbiol Immunol*, 198, 157-74.
- WHEELDON, L. J., WORTHINGTON, T., HILTON, A. C., ELLIOTT, T. S. & LAMBERT, P. A. (2008) Physical and chemical factors influencing the germination of Clostridium difficile spores. *J Appl Microbiol*, 105, 2223-30.
- WICKNER, W. (1976) Asymmetric orientation of phage M13 coat protein in Escherichia coli cytoplasmic membranes and in synthetic lipid vesicles. *Proceedings of the National Academy of Sciences*, 73, 1159-1163.
- WILLIAMSON, P. & MATTHEWS, R. (1999) Development of neutralising human recombinant antibodies to pertussis toxin. *FEMS Immunology & Medical Microbiology*, 23, 313-319.
- WILSON, K. H., SHEAGREN, J. N. & FRETER, R. (1985) Population dynamics of ingested Clostridium difficile in the gastrointestinal tract of the Syrian hamster. *J Infect Dis*, 151, 355-61.
- WINTER, G., GRIFFITHS, A. D., HAWKINS, R. E. & HOOGENBOOM, H. R. (1994) MAKING ANTIBODIES BY PHAGE DISPLAY TECHNOLOGY. *Annual Review of Immunology*, 12, 433-455.
- WINTER, G. & MILSTEIN, C. (1991) Man-made antibodies. *Nature*, 349, 293-299.
- WRIGHT, A., DRUDY, D., KYNE, L., BROWN, K. & FAIRWEATHER, N. F. (2008) Immunoreactive cell wall proteins of Clostridium difficile identified by human sera. *J Med Microbiol*, 57, 750-6.
- WRIGHT, A., WAIT, R., BEGUM, S., CROSSETT, B., NAGY, J., BROWN, K. & FAIRWEATHER, N. (2005) Proteomic analysis of cell surface proteins from Clostridium difficile. *Proteomics*, 5, 2443-52.
- WU, S., KE, A. & DOUDNA, J. A. (2007) A fast and efficient procedure to produce scFvs specific for large macromolecular complexes. *Journal of Immunological Methods*, 318, 95-101.
- WU, T. T. & KABAT, E. A. (1970) An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. *J Exp Med*, 132, 211-50.
- WÜRHN, A. & PLÜCKTHUN, A. (1999) Different Equilibrium Stability Behavior of ScFv Fragments: Identification, Classification, and Improvement by Protein Engineering. *Biochemistry*, 38, 8739-8750.
- YAMAKAWA, K., KARASAWA, T., IKOMA, S. & NAKAMURA, S. (1996) Enhancement of Clostridium difficile toxin production in biotin-limited conditions. *Journal of Medical Microbiology*, 44, 111-114.
- YAMANAKA, H., INOUE, T. & IKEDA-TANAKA, O. (1996) Chicken monoclonal antibody isolated by a phage display system. *The Journal of Immunology*, 157, 1156-1162.

- YOUNG, K. W. H., MUNRO, I. C., TAYLOR, S. L., VELDKAMP, P. & VAN DISSEL, J. T. (2007) The safety of whey protein concentrate derived from the milk of cows immunized against *Clostridium difficile*. *Regulatory Toxicology and Pharmacology*, *47*, 317-326.
- ZACCOLO, M., GRIFFITHS, A. P., PROSPERO, T. D., WINTER, G. & GHERARDI, E. (1997) Dimerization of Fab fragments enables ready screening of phage antibodies that affect hepatocyte growth factor scatter factor activity on target cells. *European Journal of Immunology*, *27*, 618-623.
- ZEILSTRA-RYALLS, J., FAYET, O. & GEORGOPOULOS, C. (1991) The universally conserved GroE (Hsp60) chaperonins. *Annual Review of Microbiology*, *45*, 301-25.
- ZHAO, B. H. & XU, C. B. (2001) [Cloning and expression of ScFv gene against alpha-toxin of *Clostridium perfringens* type A]. *Sheng Wu Gong Cheng Xue Bao*, *17*, 543-7.
- JKERSTRM, B., NILSON, B. H. K., HOOGENBOOM, H. R. & BJRCK, L. (1994) On the interaction between single chain Fv antibodies and bacterial immunoglobulin-binding proteins. *Journal of Immunological Methods*, *177*, 151-163.