

**Characterisation of the Binding Specificity and
Functional Roles of *Mycobacterium tuberculosis* Complex
CFP-10/ESAT-6 Family Proteins**

Kirsty Lynn Lightbody, MSci (Gla)

Thesis submitted for the Degree of Doctor of Philosophy
Department of Biochemistry
University of Leicester

April 2006

UMI Number: U217093

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U217093

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

Firstly I would like to thank the Veterinary Laboratories Agency for the sponsorship and generous funding which made this project possible. My thanks go to the TB Research Group headed by Prof. Glyn Hewinson, and in particular to Dr. Stephen Gordon who served on my committee, providing invaluable and enthusiastic feedback.

I would like to express thanks to my supervisor, Dr. Mark Carr, for the opportunity to work on such an interesting project. I greatly appreciate his support and the direction he has provided me throughout. My gratitude extends to those members of the group, both past and present, who have contributed to a stimulating working environment; especially to Lorna, Vaclav, Ojay and Ron. My most sincere thanks go to Dr. Philip Renshaw with whom I have worked closely on this project. His guidance, wealth of experience and friendly encouragement has been greatly appreciated.

I would like to thank my collaborators; Dr. Roger Buxton and Debbie Hunt from the National Institute of Medical Research, as well as Dr. Jim Norman, Dr. Patrick Caswell and Dr. Bernard Burke from the University of Leicester for all their help and advice.

To my friends and family, thanks for a life (and cocktails) outside of the lab; especially to Oli for his patience and support through the challenges of my project. Finally to my Mum, Dad and brothers, Liam and Gary, whose love and belief in me has supported me throughout my studies.

Kirsty Lightbody, Leicester, April 2006.

Abstract

The secreted *Mycobacterium tuberculosis* complex proteins CFP-10 and ESAT-6 form a tight, 1:1 heterodimeric complex which is known to play an essential, but as yet undefined, role in tuberculosis pathogenesis.

CFP-10 and ESAT-6 are members of a large protein family, including 23 members within the *M. tuberculosis* genome. Like CFP-10 and ESAT-6 the majority of family members are located in pairs within the genome. Yeast two-hybrid studies reported here show that the genome partners Rv0287/Rv0288 and Rv3019c/Rv3020c also form heterodimeric complexes, suggesting that all genome pairs within the CFP-10/ESAT-6 family are likely to form complexes. Further yeast two-hybrid analysis also revealed that closely related pairs, such as Rv0287/Rv0288 and Rv3019c/Rv3020c are able to bind to non-genome partners, however, distantly related proteins, including CFP-10 and ESAT-6, are unlikely to form complexes with family members other than their genome partner. The ability of some CFP-10/ESAT-6 family proteins to form complexes with non-genome partners greatly increases the number of potential complexes, and may have a significant effect on the functional diversity of this important protein family.

The recently determined structure of the CFP-10•ESAT-6 complex reveals that both proteins adopt an elongated helix-turn-helix hairpin structure and lie antiparallel to each other to form a stable four helix bundle. The surface and structural features of the complex suggest a role in binding one or more target proteins. Fluorescence microscopy studies described here have demonstrated specific binding of the CFP-10•ESAT-6 complex to the surface of monocyte and macrophage cells, suggesting the presence of a specific receptor on host cells and a possible cell role in pathogen-host cell signalling. A striking feature of the complex is the flexible C-terminal region of CFP-10, which was found to be essential for binding to host cells.

Contents

Acknowledgements.....	1
Abstract.....	2
Contents	3
Abbreviations.....	9
Chapter 1	13
Introduction	13
1.1 Tuberculosis: The Current Situation.....	13
1.2 The Organism	17
1.2.1 The <i>Mycobacterium tuberculosis</i> Complex.....	17
1.2.2 Current and Future Vaccines	20
1.2.3 Host – Microbe Interaction	22
1.3 Genomics	27
1.3.1 <i>M. tuberculosis</i> Genome	27
1.3.2 Comparative and Functional Genomics.....	28
1.4 Mycobacterial Secreted Proteins	32
1.5 CFP-10 and ESAT-6.....	34
1.5.1 The CFP-10/ESAT-6 Protein Family	34
1.5.2 Potent T-cell Antigens	37
1.5.3 Structure and Function of CFP-10 and ESAT-6.....	38
1.6 Aims of the Project	42
Chapter 2	44
Specificity of Interactions between Members of the CFP-10/ESAT-6 Family	44
2.1 Introduction.....	44
2.2 Materials and Methods.....	49
2.2.1 Strains, Plasmids and Media.....	49

2.2.2 Construction of Yeast Two-Hybrid Vectors	49
2.2.3 Transformation of <i>S. cerevisiae</i>	52
2.2.4 β -galactosidase Assays	53
2.2.4.1 Colony Filter Lift Assays.....	53
2.2.4.2 Quantitative Liquid Assays.....	53
2.3 Results.....	55
2.3.1 Cloning of CFP-10/ESAT-6 Family Members.....	55
2.3.2 Complex Formation between Genome Partners of the CFP-10/ESAT-6 Family	58
2.3.3 Interactions between Non-Genome Partners of the CFP-10/ESAT-6 Family	64
2.4 Discussion.....	68
Chapter 3	78
Specific Binding of the CFP-10•ESAT-6 Complex to the Surface of Host Cells	78
3.1 Introduction.....	78
3.2 Materials and Methods.....	81
3.2.1 Protein Expression and Purification	81
3.2.1.1 Protein Expression Vectors.....	81
3.2.1.2 Expression and Purification of CFP-10	81
3.2.1.3 Expression and Purification of ESAT-6	83
3.2.2 Expression and Purification of MPB70	83
3.2.3 Fluorescent Labelling of the CFP-10•ESAT-6 Complex and MPB70	84
3.2.4 Cell Culture.....	85
3.2.4.1 Primary Cells	85
3.2.4.2 Monocytic Cell Lines.....	85
3.2.4.3 Fibroblast Cell Lines.....	86
3.2.5 Fluorescence Microscopy	86

3.3 Results.....	88
3.3.1 Fluorescent Labelling of the CFP-10•ESAT-6 Complex and MPB70	88
3.3.1.1 Expression and Purification of CFP-10	88
3.3.1.2 Expression and Purification of ESAT-6	91
3.3.1.3 Expression and Purification of MPB70	94
3.3.1.4 Labelling of the CFP-10•ESAT-6 Complex and MPB70 with Alexa Fluor 546	96
3.3.2 Fluorescence Microscopy	96
3.3.2.1 Binding of the CFP-10•ESAT-6 Complex to Specific Host Cells	96
3.3.2.2 Specificity of CFP-10•ESAT-6 Binding to Host Cells.....	98
3.4 Discussion.....	100
Chapter 4	104
Assessment of the Functional Role of the Flexible C-termini of CFP-10 and ESAT-6..	104
4.1 Introduction.....	104
4.2 Materials and Methods.....	108
4.2.1 Cloning of C-terminal Truncated CFP-10	108
4.2.2 Expression and Purification of C-terminal Truncated CFP-10.....	108
4.2.3 Expression and Purification of C-terminal Truncated ESAT-6.....	109
4.2.4 Electrospray Mass Spectrometry Sample Preparation.....	110
4.2.5 Fluorescence Assays of Complex Formation between Variant CFP-10/ESAT- 6 Proteins	110
4.2.6 Binding of Full Length and Truncated CFP-10•ESAT-6 Complexes to Host Cells	110
4.2.6.1 Fluorescent Labelling of CFP-10•ESAT-6 Complexes	110
4.2.6.2 U937 Cell Culture	111
4.2.6.3 Fluorescence Microscopy	111

4.2.7 Calculation of the Solvent Exposure of Tryptophan Residues	111
4.3 Results.....	112
4.3.1 Cloning of C-terminal Truncated CFP-10	112
4.3.2 Expression and Purification of C-terminal Truncated CFP-10.....	113
4.3.3 Expression and Purification of C-terminal Truncated ESAT-6.....	117
4.3.4 Fluorescence Assays of Complex Formation between Variant CFP-10/ESAT-6 Proteins	121
4.3.5 Labelling of the CFP-10•ESAT-6 Complexes with Alexa Fluor 546	124
4.3.6 Fluorescence Microscopy Assays of Variant CFP-10•ESAT-6 Complexes Binding to U937 Cells	125
4.4 Discussion.....	127
Chapter 5	133
Conclusions and Future Work	133
5.1 Conclusions.....	133
5.1.1 Interactions between Members of the CFP-10/ESAT-6 Family.....	133
5.1.2 Binding of the CFP-10•ESAT-6 Complex to Specific Host Cells	134
5.1.3 Host Cell Binding of C-terminal Truncated CFP-10•ESAT-6 Complexes ..	136
5.2 Future Work.....	136
5.2.1 Identification of Host Cell Receptors for the CFP-10•ESAT-6 Complex	136
5.2.2 Determination of the Function and Mechanism of Action of the CFP-10•ESAT-6 Complex	138
5.2.3 Investigation of the Structure and Function of CFP-10/ESAT-6 Family Complexes	139
Appendix.....	140
Appendix 1 - Strains, Media and Reagents.....	140
1.1 Bacterial and Yeast Strains	140

1.2	Bacterial Culture Media.....	140
1.3	Yeast Culture Media	141
1.4	β -galactosidase Assay Solutions.....	141
Appendix 2 - Primers.....		142
Appendix 3 - Protocols.....		143
3.1	DNA Analysis by Electrophoresis	143
3.2	Restriction Digest	143
3.3	Preparation of Competent Cells.....	144
3.4	Transformation of Competent Cells	145
3.5	SDS-Polyacrylamide Gel Electrophoresis	145
Appendix 4 - DNA Sequencing.....		148
4.1	pGBD/CFP-10 Chromatogram	148
4.2	pGAD/CFP-10 Chromatogram.....	149
4.3	pGBD/ESAT-6 Chromatogram	150
4.4	pGAD/ESAT-6 Chromatogram	151
4.5	pGBD/Rv0287 Chromatogram.....	152
4.6	pGAD/Rv0287 Chromatogram.....	153
4.7	pGBD/Rv0288Chromatogram	154
4.8	pGAD/Rv0288 Chromatogram.....	155
4.9	pGBD/Rv3019c Chromatogram	156
4.10	pGAD/Rv3019c Chromatogram.....	157
4.11	pGBD/Rv3020c Chromatogram	158
4.12	pGAD/Rv3020cChromatogram	159
4.13	pET28a/tCFP-10 Chromatogram.....	160
References.....		161

Published Work 177

Lightbody, K.L., Renshaw, P.S., Collins, M.L., Wright, R.L., Hunt, D.M., Gordon, S.V., Hewinson, R.G., Buxton, R.S., Williamson, R.A. and Carr, M.D. (2004) Characterisation of complex formation between members of the *Mycobacterium tuberculosis* complex CFP-10/ESAT-6 protein family: towards an understanding of the rules governing complex formation and thereby functional flexibility. *FEMS Microbiol. Lett.* 238, 255-262.

Renshaw, P.S., Lightbody, K.L., Veverka, V., Muskett, F.W., Kelly, G., Frenkiel, T.A., Gordon, S.V., Hewinson, R.G., Burke, B., Norman, J., Williamson, R.A. and Carr, M.D. (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J.* 24, 2491-2498.

Abbreviations

Single and Three Letter Codes for Amino Acids

Alanine	A	Ala	Leucine	L	Leu
Arginine	R	Arg	Lysine	K	Lys
Asparagine	N	Asn	Methionine	M	Met
Aspartic acid	D	Asp	Phenylalanine	F	Phe
Cysteine	C	Cys	Proline	P	Pro
Glutamine	Q	Gln	Serine	S	Ser
Glutamic Acid	E	Glu	Threonine	T	Thr
Glycine	G	Gly	Tryptophan	W	Trp
Histidine	H	His	Tyrosine	Y	Tyr
Isoleucine	I	Ile	Valine	V	Val

DNA Bases

Adenine	A	Guanine	G
Cytosine	C	Thymine	T

Abbreviations

-leu/-his/-trp	minus leucine/minus histidine/minus tryptophan
3-AT	3-amino-1,2,4-triazole
ADH1	Alcohol dehydrogenase 1
Ag85	Antigen 85
AIDS	Acquired immunodeficiency syndrome
Amp	Ampicillin
APC	Antigen presenting cell
ATP	Adenosine-5'-triphosphate
BAC	Bacterial artificial chromosome
B-cell	Bursa dependant lymphocyte
BCG	Bacille Calmette-Guerin
Bis-Tris	2-[bis(2-hydroxyethyl)amino]-2-hydroxymethyl-propane-1,3-diol
BSA	Bovine serum albumin
bp	Base pair
CFP-10	Culture filtrate protein of 10 kD
CFP-10/ESAT-6	Refers to CFP-10 and ESAT-6 as two proteins

CFP-10•ESAT-6	Refers to the 1:1 complex
Da	Dalton
DBD	DNA binding domain
dH ₂ O	De-ionised water
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DOTS	Directly observed therapy short-course
DTH	Delayed type hypersensitivity
DTT	1,4-dithiothreitol
EDTA	Ethylene diaminetetraacetate
ESAT-6	Early secreted antigenic target of 6 kD
FCS	Foetal calf serum
Gdn-HCl	Guanidine hydrochloride
HIV	Human immunodeficiency virus
IFN γ	Interferon gamma
IGP	Imidazole glycerol phosphate
IL	Interleukin
INH	Isoniazid
IPTG	Isopropyl-1-thio- β -D-galactoside
Kan	Kanamycin
kDa	kilo Dalton
LB	Luria-Bertani
LDS	Lithium dodecyl sulphate
LiAc	Lithium acetate
MCS	Multiple cloning site
MDR	Multi-drug resistant
MES	2-(N-morpholino)ethanesulphonic acid
MM6	Monocyte macrophage 6
MPB	<i>Mycobacterium bovis</i> protein
MPT	<i>Mycobacterium tuberculosis</i> protein
MW	Molecular weight
ms	Milli-seconds
OD	Optical density
ONPG	<i>ortho</i> -nitrophenyl- β -D-galactopyranoside
ONP	<i>ortho</i> -nitrophenol

ORF	Open reading frame
P	Promoter
PAGE	Polyacrilamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PE	Proline-Glutamic acid sequence
PFA	Paraformaldehyde
<i>Pfu</i>	<i>Pyrococcus furiosus</i>
PMSF	Phenylmethylsulphonyl fluoride
PPE	Proline-Proline-Glutamic acid sequence
RD	Region of difference
RIF	Rifampicin
RPM	Revolutions per minute
SD	Synthetic dropout
SDS	Sodium dodecyl sulphate
ST-CF	Short term-culture filtrate
T	Terminator
TAD	Transcription activation domain
TAE	Tris-acetate
TB	Tuberculosis
T-cell	Thymus dependant lymphocyte
tCFP-10	Truncated CFP-10
tESAT-6	Truncated ESAT-6
TE	Tris HCl-EDTA buffer
TNF α	Tumor necrosis factor alpha
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
TST	Tuberculin skin test
UAS	Upstream activating sequence
UV	Ultraviolet
v/v	Volume per volume
w/v	Weight per volume
WHO	World Health Organisation
X-gal	5-bromo-4-chloro-3-indolyl- β -D-galactoside
YPD	Yeast extract, peptone, dextrose

Table of the Rv numbers and protein names for members of the *Mycobacterium tuberculosis* CFP-10/ESAT-6 family. Throughout this thesis CFP-10 and ESAT-6 will be referred to as CFP-10 and ESAT-6 and other members of the family will primarily be referred to using the Rv number.

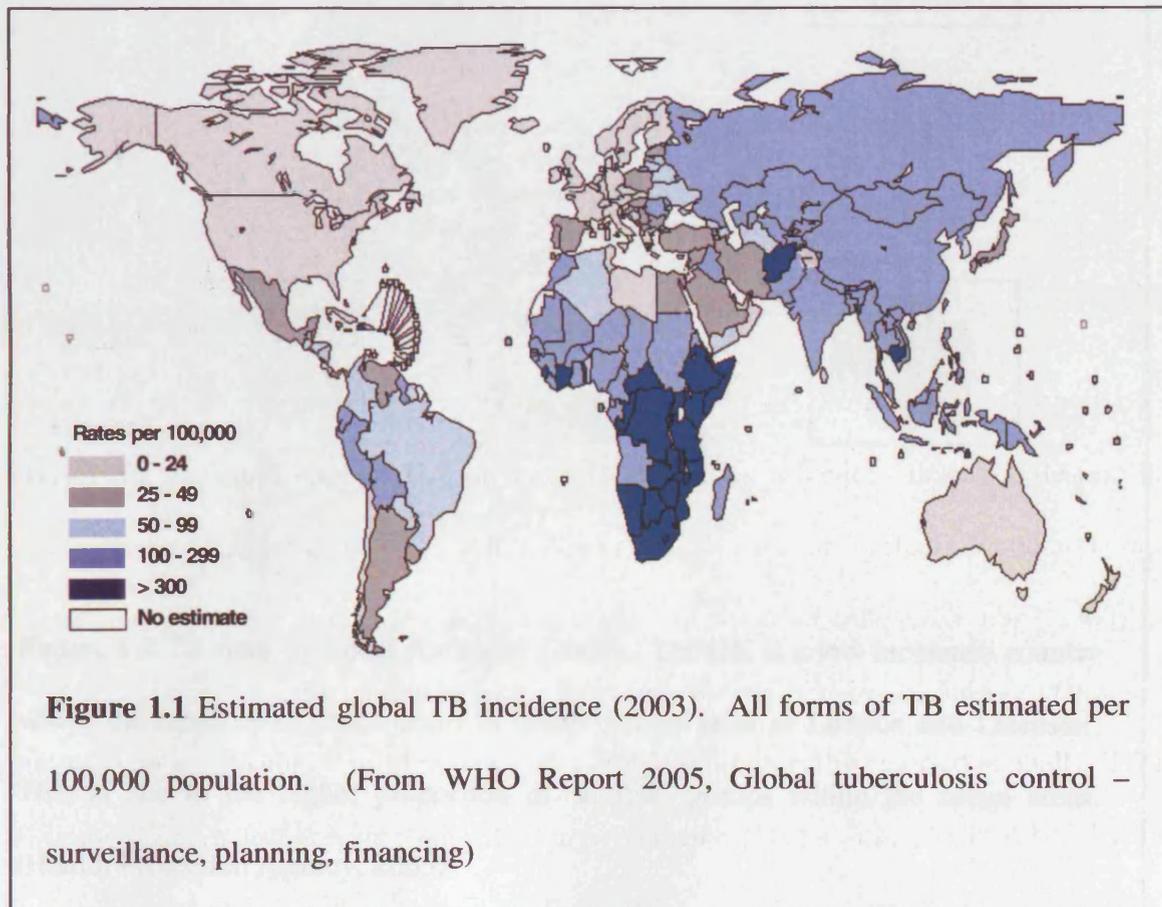
Rv Number	Name
Rv3874	esxB, CFP-10
Rv3875	esxA, ESAT-6
Rv3890c	esxC
Rv3891c	esxD
Rv3904c	esxE
Rv3905c	esxF
Rv0287	esxG, TB 9.8
Rv0288	esxH, TB10.4, CFP-7
Rv1037c	esxI, TB 9.9D (MTINY Family)
Rv1038c	esxJ, TB 11.0, (QILLS Family)
Rv1197	esxK, TB11.0, (QILLS Family)
Rv1198	esxL, TB 9.9C, (MTINY Family)
Rv1792	esxM, TB 11.0, (QILLS Family)
Rv1793	esxN, TB 9.9A, (MTINY Family)
Rv2346c	esxO, TB 9.9E, (MTINY Family)
Rv2347c	esxP, (QILLS Family)
Rv3017c	esxQ, TB 12.9
Rv3019c	esxR, TB 10.3
Rv3020c	esxS
Rv3444c	esxT
Rv3445c	esxU
Rv3619c	esxV, TB 9.9D (MTINY Family)
Rv3620c	esxW (QILLS Family)

Chapter 1

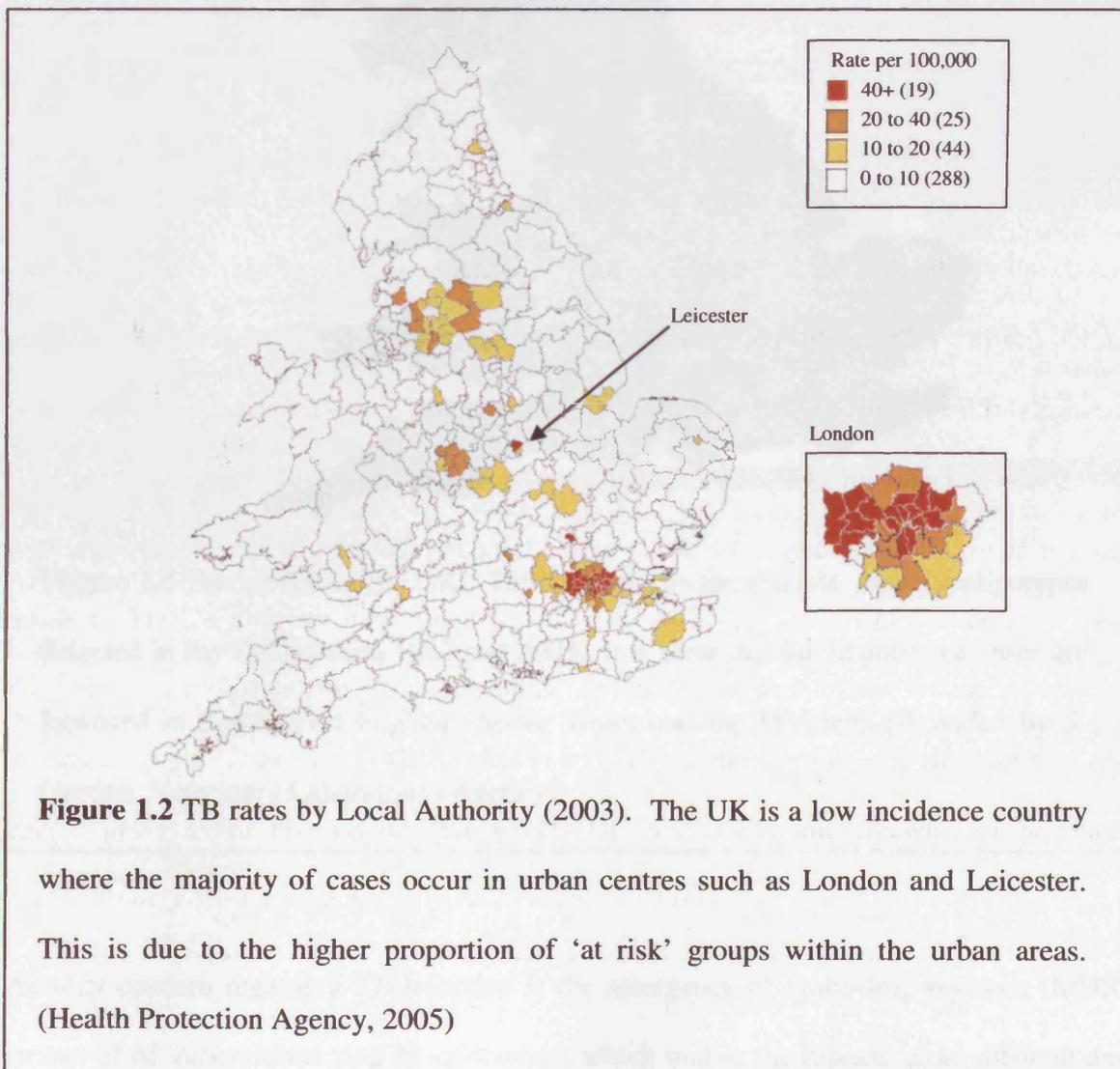
Introduction

1.1 Tuberculosis: The Current Situation

Tuberculosis (TB) was declared a global emergency by the World Health Organisation (WHO) in 1993 and remains one of the world's leading infectious diseases, despite the availability of the BCG (Bacille Calmette-Guerin) vaccine and effective chemotherapy (Raviglione, 2003). Currently, two billion people (one third of the world's population) are infected with TB bacilli, predominantly *Mycobacterium tuberculosis* and *M. africanum*, and with approximately 9 million new cases and 2 million deaths reported annually, TB is a major concern for human, and veterinary, medicine (Dye *et al.*, 1999, WHO Report 2005).



Almost all human TB cases (95 %) and deaths (98 %) occur in developing countries, particularly in South-East Asia and Africa, where the problem is exacerbated by the spread of HIV/AIDS (Figure 1.1). The WHO reported that by 2001 almost 11.5 million people were co-infected with HIV and TB, of which over 70 % of cases occurred in Sub-Saharan Africa (WHO Report 2004a). Britain is considered to be a low incidence country with respect to TB infection, however reported cases have increased by 20 % over the last decade. The majority of cases in the UK occur in urban areas where there is a high proportion of high risk sub-groups, such as migrants from high burden countries and the homeless (Figure 1.2). In 2002, London accounted for almost 44 % of cases in England and Wales (Health Protection Agency, TB Update, 2005), and likewise in Scotland, approximately 50 % of reported TB cases occur in Glasgow.



In contrast to human TB infection, the disease in cattle, caused by *M. bovis*, is a particular problem in the UK. Most cases of bovine TB are clustered in South-West England, South Wales and the West Midlands (Figure 1.3), accounting for almost 89 % of confirmed new incidents (The Report of the Chief Veterinary Officer, 2005). However, restocking of cattle after the foot and mouth outbreak (2001) has led to increased incidence in northern England and Scotland.

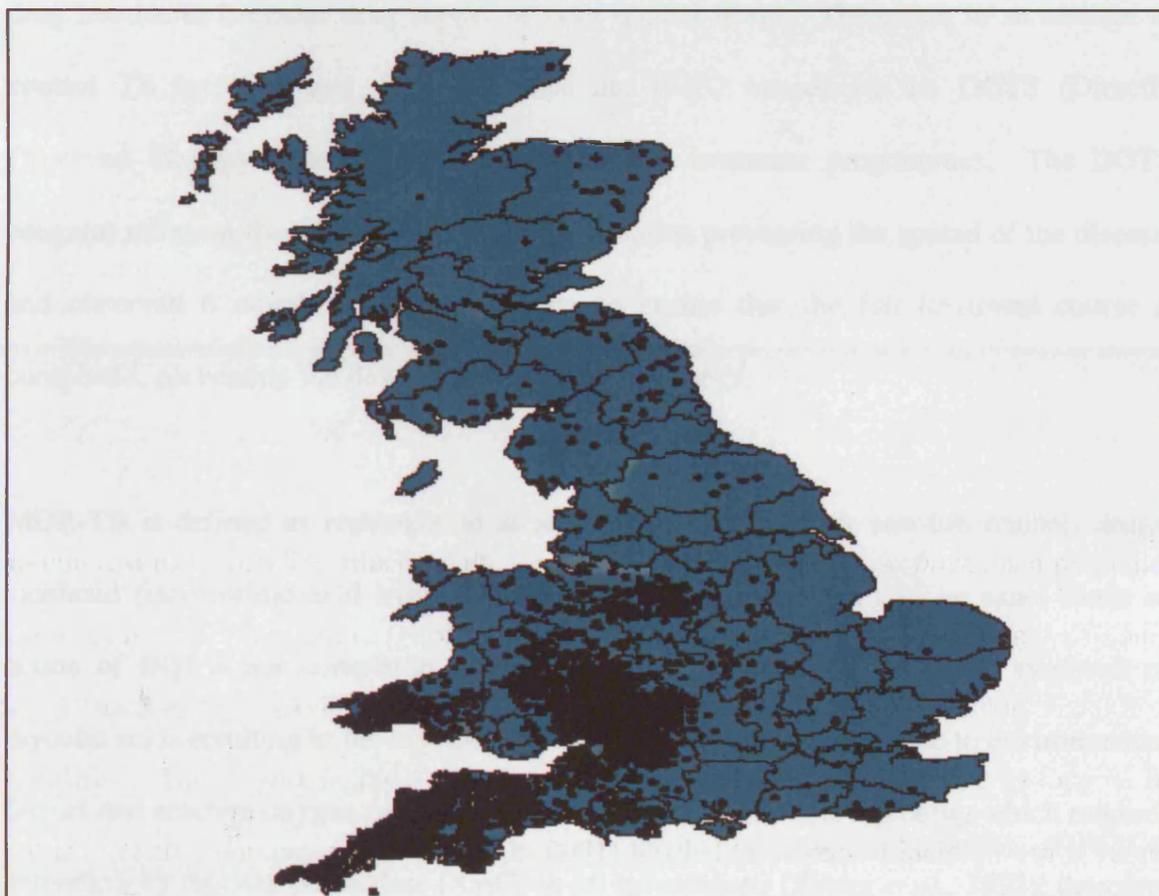


Figure 1.3 Bovine TB in the UK. This map shows the total *M. bovis* spoligotypes detected in the UK between 1987 and 2003. It is clear that the majority of cases are localised in South-West England, South Wales and the Midlands (Provided by S. Gordon, Veterinary Laboratories Agency)

Another concern regarding TB infection is the emergence of multi-drug resistant (MDR) strains of *M. tuberculosis* and *M. africanum*, which makes the disease more difficult and

more expensive to treat. Globally it is estimated that 50 million people are infected with MDR strains of mycobacteria, and approximately 300,000 new MDR TB cases were reported in 2003 (Kaufmann, 2005). Infection with MDR TB is highly prevalent in countries including the Russian Federation, China (Hanan and Liaoning Provinces) and former Soviet Union countries (Kazakhstan, Estonia, Latvia, Lithuania) (WHO Report, 2004b). Drug resistance in TB infection often develops as a result of poor adherence to drug treatment, irregular drug supply or poor quality drugs. Therefore, in an attempt to control TB infection and drug resistance the WHO introduced the DOTS (Directly Observed Therapy Short-Course) and DOTS-plus treatment programmes. The DOTS program relies on the prompt detection of new cases, preventing the spread of the disease, and observed 6 month treatment regimens to ensure that the full treatment course is completed, preventing the development of drug resistance.

MDR-TB is defined as resistance to at least two of the primary anti-tuberculosis drugs, isoniazid (isonicotinic acid hydrazide, INH) and rifampicin (RIF). The exact mode of action of INH is not completely understood, but is thought to inhibit the synthesis of mycolic acids resulting in the mycobacterium becoming more susceptible to environmental factors and reactive oxygen radicals (Rattan *et al.*, 1998). INH is a prodrug which requires activation by catalase-peroxidase (KatG) in *M. tuberculosis* (Zhang *et al.*, 1992); therefore mutations which abolish or alter the activity of KatG are often associated with INH resistance. A common mutation conferring resistance to INH is an amino acid substitution (serine to threonine) at residue 315 in *katG*, which results in decreased catalase activity and a reduced ability to activate INH (Musser *et al.*, 1996, Ramaswamy and Musser, 1998). INH resistance has also been associated with the *mabA-inhA* operon, where mutations within the *inhA* gene prevent INH inhibition of mycolic acid biosynthesis and mutations upstream of the *mabA-inhA* operon lead to over-expression of the drug target (InhA)

(Banerjee *et al.*, 1994, Ramaswamy and Musser, 1998). Mutations relating to INH resistance have also been reported in *kasA*, part of a six gene operon involved in mycolic acid synthesis (Mdluli *et al.*, 1998, Slayden and Barry, 2000). The mechanism of action of RIF is to inhibit mycobacterial transcription by binding to the β -subunit of RNA polymerase. Mutations within the *rpoB* gene, which encodes the β -subunit of RNA polymerase, are known to result in resistance to RIF. Over 95% of RIF resistant isolates possess mutations, including point mutations, short deletions and insertions, within an 81-bp rifampicin-resistance determining region (RRDR) of *rpoB*, resulting in alterations of the structure of RpoB (Telenti *et al.*, 1993, Ramaswamy and Musser, 1998).

1.2 The Organism

1.2.1 The *Mycobacterium tuberculosis* Complex

The primary agents of human TB are *M. tuberculosis* and *M. africanum* which are Gram positive, acid-fast bacilli (Figure 1.4). *M. tuberculosis* and *M. africanum* along with four other members of the slow growing mycobacteria form the *M. tuberculosis* complex, a group of closely related organisms characterised by a high level of genetic similarity (Sreevatsan *et al.*, 1997b, Musser *et al.*, 2000). The complex includes; *M. africanum* (Frothingham *et al.*, 1999), the prominent cause of TB in Africa, *M. bovis*, an important veterinary pathogen capable of infecting a wide range of hosts including cattle and humans (De la Rua-Domenech, 2005), *M. bovis* BCG (Bacille-Calmette Guerin), the laboratory attenuated *M. bovis* strain commonly used as a live vaccine against TB, *M. canettii*, a rare agent of human infection (Van Soolingen *et al.*, 1997), *M. microti*, which typically infects rodents such as mice and voles and is very rarely associated with human disease (Van Soolingen *et al.*, 1998, Nineman *et al.*, 2000, Frota *et al.*, 2004) and *M. tuberculosis* (Figure 1.5). The members of the complex are considered to be a single species (Imaeda, 1985, Feizabadi *et al.*, 1996, Frothingham *et al.*, 1994) originating from a *M. tuberculosis*

progenitor strain (Brosch *et al.*, 2002). Despite high levels of genetic similarity between the *M. tuberculosis* complex the strains differ in pathogenicity, geography and host range (Smith *et al.*, 2005).

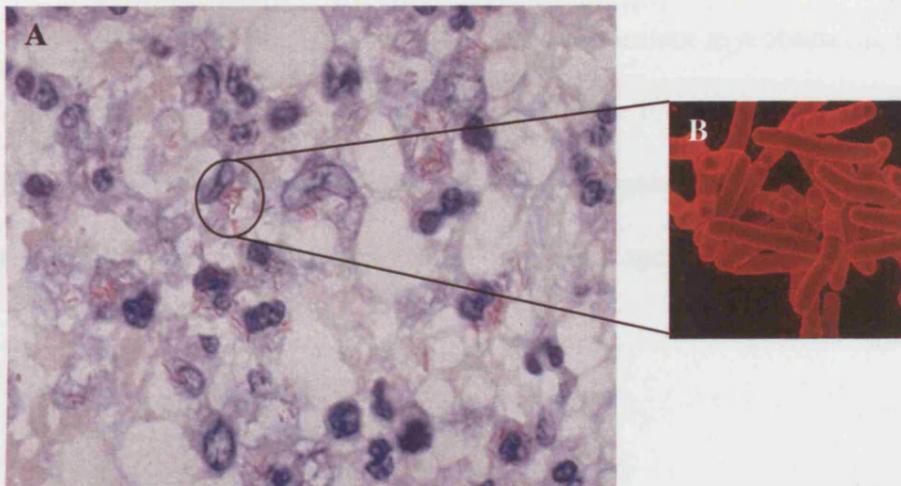


Figure 1.4 *M. tuberculosis* is an acid-fast, Gram positive, rod shaped bacillus. (A) Ziehl-Neelsen stained *M. tuberculosis* bacilli (red) in infected lung tissue (<http://www-medlib.med.utah.edu/WebPath/HISTHTML/STAINS/STAIN017.html>). (B) Scanning electron micrograph of *M. tuberculosis*, adapted from the image posted at the National Institute of Health, Laboratory of Immunogenetics, TB Research Section (www.niaid.nih.gov/dir/labs/lhd/barry.htm).

Both *M. bovis* BCG and *M. microti* have been used with equal success as vaccines against human TB (Hart and Sutherland, 1977). Serial passage of virulent *M. bovis* on potato-glycerine-ox bile agar between 1908 and 1921 gave rise to a strain that retained immunogenicity but demonstrated decreased virulence in animals and humans. Unable to preserve the original BCG strain, the vaccine was produced by serial passage until the introduction of lyophilisation in the 1960's, which resulted in a number of genetically different daughter strains. The molecular basis for attenuation is yet to be fully understood, however all BCG sub-strains lack a small portion of the *M. bovis* genome

known as region of difference 1 (RD1). RD1 includes the genes encoding CFP-10 (Rv3874 or *esxB*) and ESAT-6 (Rv3875 or *esxA*) (Mahairas *et al.*, 1996, Behr *et al.*, 1999), and is considered to be the original attenuating mutation of *M. bovis* BCG (Pym *et al.*, 2002). Interestingly, *M. microti* naturally lacks the genes encoding CFP-10 and ESAT-6 (Brodin *et al.*, 2002). The presence of these genes in virulent mycobacteria, including *M. tuberculosis*, *M. bovis* and *M. leprae*, and their absence in the attenuated *M. bovis* BCG and *M. microti* bacteria indicate a clear role in pathogenesis and have also led to the inclusion of CFP-10 and ESAT-6 as potential diagnostic agents.

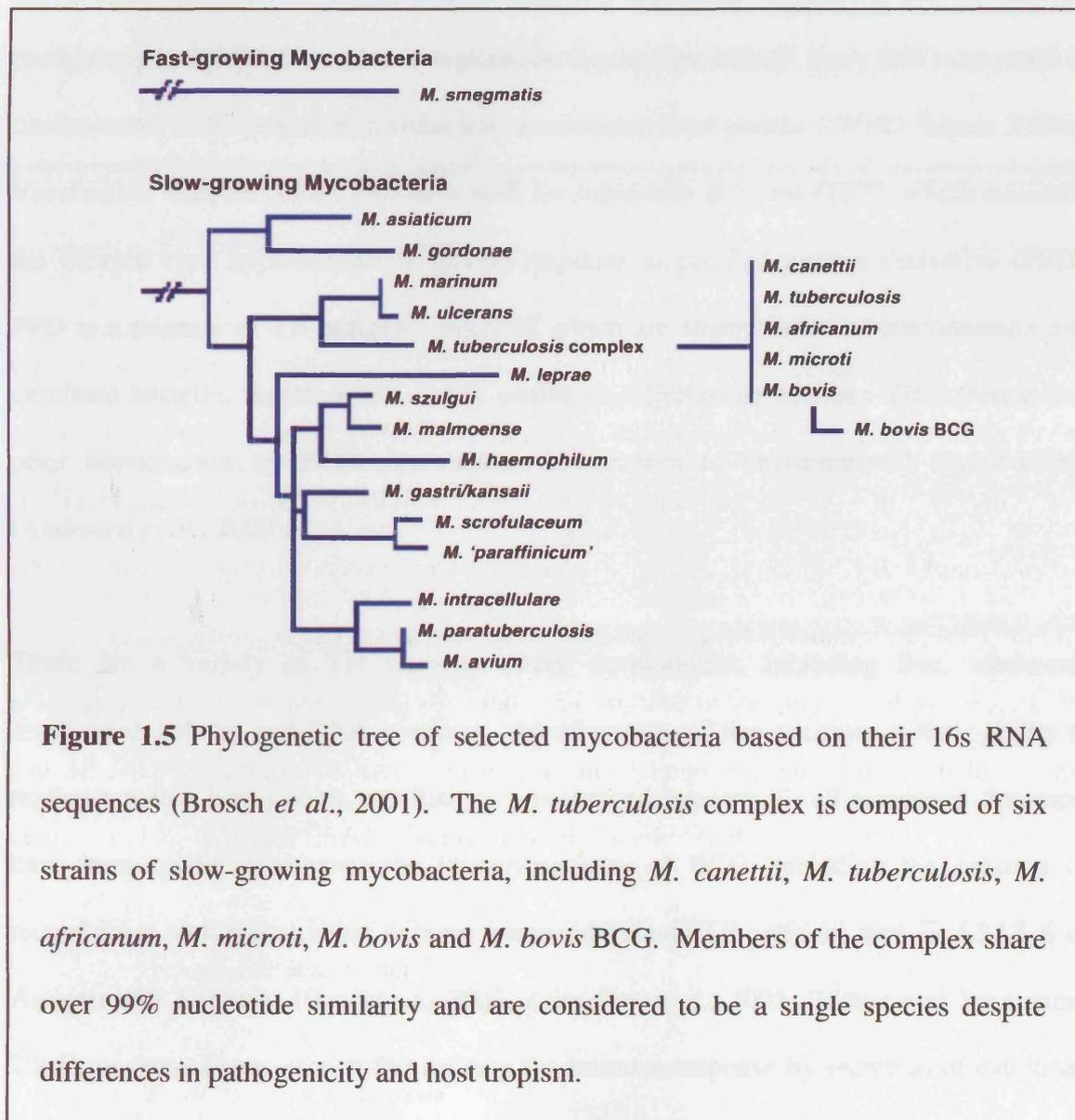


Figure 1.5 Phylogenetic tree of selected mycobacteria based on their 16S RNA sequences (Brosch *et al.*, 2001). The *M. tuberculosis* complex is composed of six strains of slow-growing mycobacteria, including *M. canettii*, *M. tuberculosis*, *M. africanum*, *M. microti*, *M. bovis* and *M. bovis* BCG. Members of the complex share over 99% nucleotide similarity and are considered to be a single species despite differences in pathogenicity and host tropism.

1.2.2 Current and Future Vaccines

M. bovis BCG is currently used as a vaccine to protect against TB. It is estimated that approximately 1 billion people have received BCG, however, studies have suggested that the efficacy this vaccine can range from 0-80 %. The exact reason for this variability is not known, although human genetics, the use of various BCG daughter strains or prior exposure to environmental mycobacteria may play a role (Fine *et al.*, 1999, Brandt *et al.*, 2002). It has also been proposed that serial passage of *M. bovis* BCG in the lab may have led to over-attenuation and reduced protection against TB (Behr and Small, 1997, Mostowy *et al.*, 2003). The safety of this live attenuated vaccine is also a concern, particularly in HIV/AIDS endemic regions, as vaccination with *M. bovis* BCG can result in disseminated BCG infection in some immunocompromised patients (WHO Report 2004a). Vaccination with BCG also interferes with the tuberculin skin test (TST), which measures the delayed type hypersensitivity (DTH) response to purified protein derivative (PPD). PPD is a mixture of TB antigens, many of which are shared between mycobacteria and unrelated bacteria, therefore this test is unable to differentiate between TB infection and prior sensitisation by BCG vaccination or exposure to environmental mycobacteria (Andersen *et al.*, 2000).

There are a variety of TB vaccines under development, including live, whole-cell inactivated, subunit and DNA vaccines. An advantage of live vaccines is their ability to replicate within host tissues, resulting in a prolonged memory T-cell response. Attempts have been made to enhance the immunogenicity of BCG; including the creation of recombinant strains that either express immunodominant TB antigens such as ESAT-6 or Antigen 85B (Ag85B) (Bao *et al.*, 2003, Palendira *et al.*, 2005, Eddine and Kaufmann, 2005), or recombinant strains that modify the immune response by secretion of cytokines or listeriolysin from *Listeria monocytogenes* (Murray *et al.*, 1996, Grode *et al.*, 2005).

However, as described previously, the use of live *M. bovis* BCG in immunocompromised patients poses a problem. Studies have shown that deletion of genes involved in amino acid biosynthesis, such as *leuD*, result in further attenuation of *M. bovis* BCG and a reduced ability to survive within macrophages (Bange *et al.*, 1996), immunocompetent mice (C57BL/6) (McAdam *et al.*, 1995) and SCID (severe combined immunodeficiency disease) mice (Guleria *et al.*, 1996). Reports have also indicated that BCG auxotrophs can confer some protection against challenge with *M. tuberculosis* or *M. bovis* in mice and guinea pigs (Guleria *et al.*, 1996, Chambers *et al.*, 2000), suggesting the *M. bovis* BCG auxotrophic mutants may represent potential vaccines suitable for use in immunocompromised individuals. In contrast to enhancing and/or attenuating BCG, rational attenuation of *M. tuberculosis* or *M. bovis* could provide a greater antigenic repertoire than BCG and may also prove safe for use in immunocompromised persons (Smith *et al.*, 2001, Sampson *et al.*, 2004, Sambandamurthy *et al.*, 2005).

An alternative live vaccine is *M. microti*. *M. microti* has been widely used as a vaccine against TB in the past and has proven to be equally as safe and effective as BCG. Antigenicity would be maintained by serial passage through voles, however live *M. microti* would pose the same problem as *M. bovis* BCG in immunocompromised individuals. Apart from live attenuated mycobacteria, subunit vaccines using potent TB antigens such as ESAT-6 and Ag85B have shown promising results in animal models (Olsen *et al.*, 2001, Olsen *et al.*, 2004, Langermans *et al.*, 2004). Another alternative is DNA vaccines, where the protein of interest is made within the host cell following vaccination with a plasmid containing a gene of interest. DNA vaccines are cheap, easy to produce and store and induce both CD4⁺ and CD8⁺ T-cell responses, however, concerns include integration of the plasmid DNA into the host genome and the potential development of auto-immune disease (Huygen *et al.*, 2005). With regard to TB DNA vaccines some studies have shown that

DNA vaccination alone confers no protection against TB challenge (Chambers *et al.*, 2002, Wedlock *et al.*, 2003). DNA vaccines could be improved by expression of multiple antigens (Kamath *et al.*, 1999, Derrick *et al.*, 2004a), by electroporation (Tollefsen *et al.*, 2003, Li *et al.*, 2005) or use in prime boost strategies along with BCG (Huygen, 2003, Skinner *et al.*, 2003, Derrick *et al.*, 2004b, Hogarth *et al.*, 2006).

1.2.3 Host – Microbe Interaction

Tuberculosis is a contagious disease, transmitted via the respiratory route. Only a low infective dose is required to initiate infection, however few infections progress to clinical disease (5-10 % in non-immunocompromised patients) (Figure 1.6). The majority of inhaled *M. tuberculosis* bacilli remain in the upper respiratory tract and are expelled by the mucociliary escalator. The organisms that do reach the lung are phagocytosed by alveolar macrophages, where the host cell either kills the bacilli eliminating the infection, or by failing to mount an effective response, allows the bacterium to establish an infection (Fenton and Vermeulen, 1996). Bacteria are also taken up by dendritic cells, which become activated and migrate to the lymph nodes to stimulate naïve T-cells. *M. tuberculosis* has evolved to survive and replicate within the macrophage phagosome, in part due to the complex lipid rich cell wall and also due to inhibition of normal phagosome maturation along the endosomal-lysosomal pathway. Early endosomes interact with phagosomes containing *M. tuberculosis*, however, fusion with late-endosomes and lysosomes is inhibited (Clemens and Horowitz, 1995, Clemens and Horowitz, 1996, Via *et al.*, 1997). The mycobacterial lipids phosphatidylinositol mannoside (PIM) and lipoarabinomannan (LAM) appear to be involved in the inhibition of phagosome maturation, where PIM stimulates fusion with early endosomes and LAM interferes with cell signalling which prevents the acquisition of late endosomal or lysosomal constituents (Fratti *et al.*, 2001, Fratti *et al.*, 2003, Vergne *et al.*, 2004).

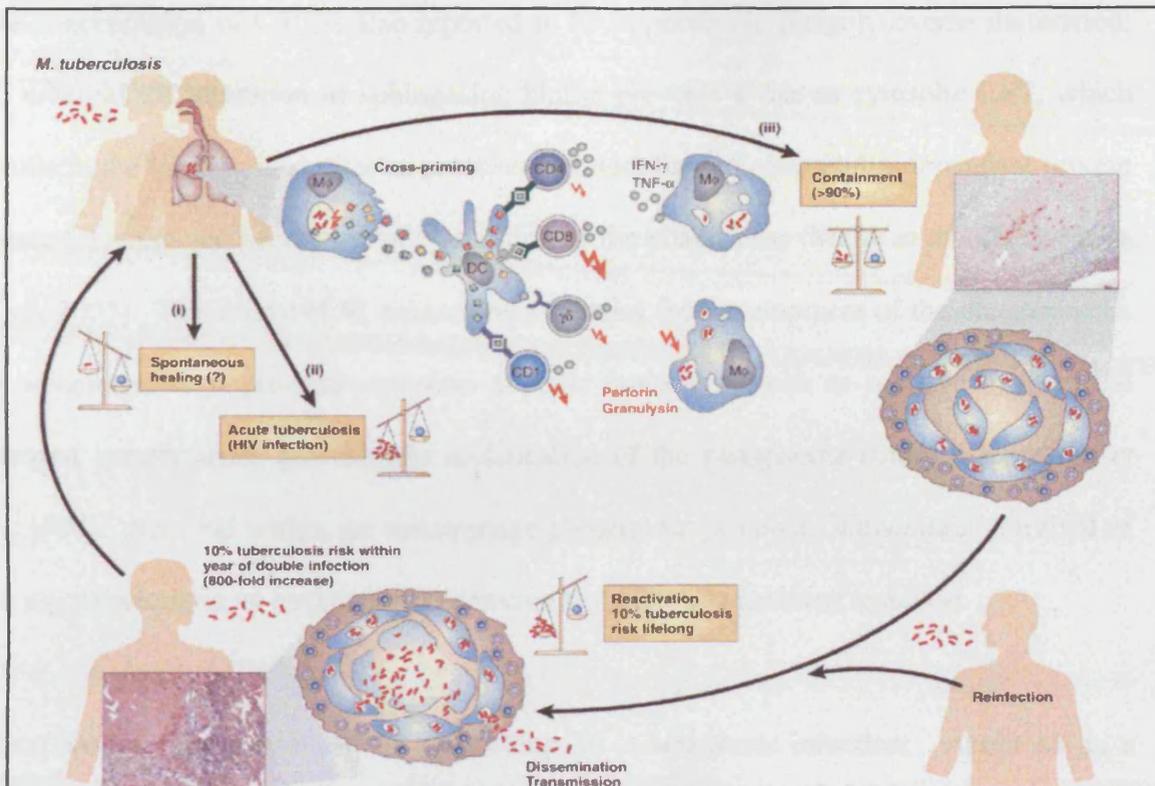


Figure 1.6 Possible outcomes of *M. tuberculosis* infection and the underlying immune mechanisms (taken from Kaufmann and McMichael, 2005). *M. tuberculosis* is transmitted via aerosol droplets (top left). There are three possible outcomes; (i) The infection is eradicated by the immune system immediately, (ii) infection progresses into tuberculosis (most common in immunocompromised individuals) and (iii) the infection is contained by the immune system. Following inhalation by the host, *M. tuberculosis* is taken up by alveolar macrophages and dendritic cells. Migration of infected dendritic cells to the lymph nodes stimulates naïve T-cell populations, including CD4⁺ and CD8⁺ T-cells. Antigenic T-cells induce granuloma formation, in which the *M. tuberculosis* infection is contained. The granuloma can contain the tubercle bacilli for years, however immunocompromise of the host (HIV, old age, malnutrition) can lead to reactivation of the infection. Dissemination of the mycobacteria throughout the lung gives the bacilli access to the respiratory system, allowing transmission to new hosts.

The concentration of Ca^{2+} is also reported to be important in phagolysosome maturation. *M. tuberculosis* inhibition of sphingosine kinase prevents a rise in cytosolic Ca^{2+} , which results in the inhibition of effector proteins (calmodulin and calmodulin-dependant protein kinase II) which are necessary for maturation of the phagosome (Malik *et al.*, 2000, Malik *et al.*, 2003). The ability of *M. tuberculosis* to arrest the development of the phagosome is an advantage, as it prevents exposure to toxic molecules, such as reactive oxygen and nitrogen intermediates and inhibits acidification of the phagosome (Sturgill-Koszycki *et al.*, 1994). Survival within the macrophage phagosome promotes intracellular survival of the mycobacteria in an environment protected from the host immune response.

Macrophages play a key role in controlling *M. tuberculosis* infection. Apoptosis is a common response to infection and is often viewed as an essential part of the host defence, preventing the spread of mycobacteria and inducing a cell mediated response by activation of T-cells (Placido *et al.*, 1997). Studies have shown that virulent mycobacteria (*M. tuberculosis* H37Rv) are less potent inducers of apoptosis than attenuated strains (*M. bovis* BCG and *M. tuberculosis* H37Ra) (Keane *et al.*, 2000, Zhang *et al.*, 2005), linking protection of mycobacteria with suppressed apoptosis. Cell necrosis, promoting bacterial dissemination and tissue damage may be more beneficial to *M. tuberculosis* (Gil *et al.*, 2004). Apoptosis and necrosis of host cells as a result of *M. tuberculosis* infection has yet to be fully understood, however, the balance of cytokines such as tumor necrosis factor α (TNF α) and interleukin-10 (IL-10) play an important role (Rojas *et al.*, 1999).

Following initial infection, the bacteria multiply within the macrophage resulting in cell lysis, allowing the bacilli to disseminate and infect additional host cells, and leads to the release of pro-inflammatory cytokines (IL-1, IL-6, IL-12 and TNF α) that initiate a cell mediated response. Macrophages and lymphocytes (B and T-cells) migrate to the site of

infection forming a granuloma (or tubercule). The production of interferon gamma (IFN γ) by CD4⁺ and CD8⁺ T-cells activates macrophages and enhances their phagocytic and microbicidal potential (Stenger and Modlin, 1999). CD8⁺ T-cells produce perforin and granulysin which lyse *M. tuberculosis* infected macrophages, killing the intracellular bacteria (Stenger *et al.*, 1998) (Figure 1.6).

Infected macrophages are surrounded by antigen specific CD4⁺ and CD8⁺ T-cells, which form the structure of the granuloma. Typically, the granuloma is associated with containment of the infection and the necrotic centre, with acidic pH, low oxygen availability and presence of toxic molecules is considered to be a hostile environment which limits replication of *M. tuberculosis*. The necrotic centre is often devoid of intact host cells, and despite the presence of mycobacterial material (DNA), few mycobacteria are detected in this region (Fenhalls *et al.*, 2002a, Fenhalls *et al.*, 2002b). A recent study by Ulrichs *et al.* (2004) has shown that the central necrotic region is surrounded by an inner layer of CD4⁺ T-cells and monocyte-derived macrophages, some of which differentiate into epithelioid cells or fuse to form giant cells. Interestingly, the study also indicated that the outer layer of the granuloma is mainly comprised of CD8⁺ T-cells which is surrounded by a fibrotic rim. The fact that these cells are not detected in the inner layer contradicts the hypothesis that CD8⁺ T-cells are involved in cytolysis of infected APCs and growth of the necrotic centre.

It has been generally accepted that the centre of the granuloma plays an important role in host-pathogen interaction. However, the study by Ulrichs *et al.* (2004) also reported the formation of follicle-like structures in the peripheral rim of the granuloma (Figure 1.7). These follicles are composed of CD4⁺ and CD8⁺ T-cells, B-cells and antigen presenting cells (APCs) which resemble secondary lymphoid organs. In addition, lymphocyte

proliferation and cytokine production primarily occurs outside of the granuloma, which supports the suggestion that these follicular structures are the primary site of host-pathogen cross-talk (Fenhalls *et al.*, 2000, Ulrichs *et al.*, 2004). In particular, TNF α is thought to play an important role in formation and maintenance of the granuloma structure. Studies in mice have shown that absence of TNF α or the TNF receptor p55 (TNFRp55) leads to abnormal granuloma formation and a reduced ability to control mycobacterial infection (Flynn *et al.*, 1995, Bean *et al.*, 1999). It is thought that TNF α regulates expression of chemokines and as a result affects migration and localisation of cells to the site of mycobacterial infection (Algood *et al.*, 2005).

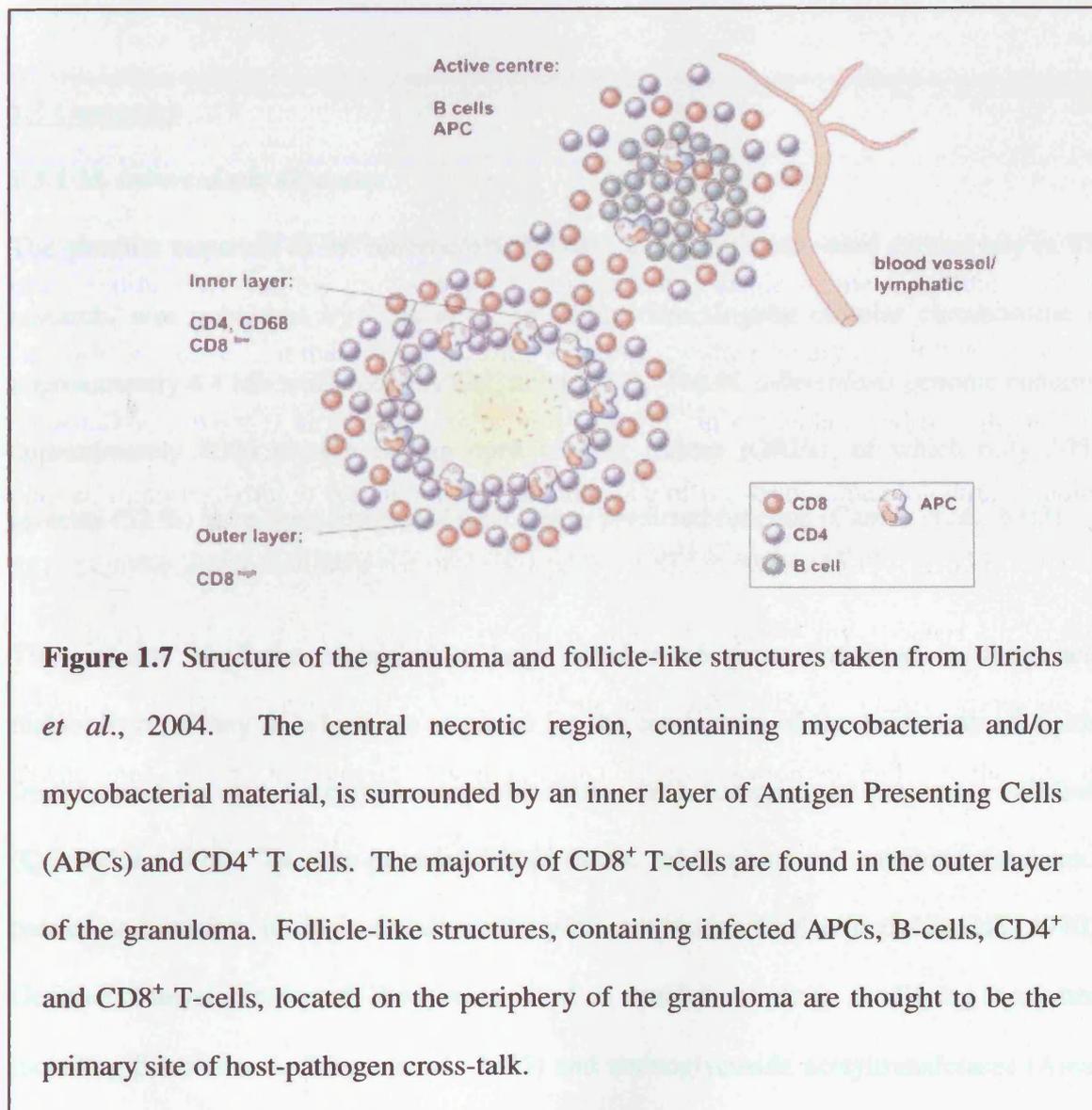


Figure 1.7 Structure of the granuloma and follicle-like structures taken from Ulrichs *et al.*, 2004. The central necrotic region, containing mycobacteria and/or mycobacterial material, is surrounded by an inner layer of Antigen Presenting Cells (APCs) and CD4⁺ T cells. The majority of CD8⁺ T-cells are found in the outer layer of the granuloma. Follicle-like structures, containing infected APCs, B-cells, CD4⁺ and CD8⁺ T-cells, located on the periphery of the granuloma are thought to be the primary site of host-pathogen cross-talk.

The strength of the cell mediated response determines whether the infection progresses to primary TB or whether latent infection develops. Latency is a delicate balance between the host immune system and the pathogen where the host is non-contagious and asymptomatic (Flynn and Chan, 2001). Post-primary TB can occur as a result of re-activation of a latent infection, often due to suppression of the host immune system (including HIV infection, substance abuse, malnutrition, stress, aging or immunosuppressive drug treatment) or re-infection following a previous infection. Uncontrolled by the immune system the bacilli disseminate throughout the lung and progress to other tissues, leading to a severe disease state. Rupturing of the bronchi gives *M. tuberculosis* access to the airway, allowing transmission to new hosts.

1.3 Genomics

1.3.1 *M. tuberculosis* Genome

The genome sequence of *M. tuberculosis* H37Rv, a virulent strain used extensively in TB research, was published by Cole *et al.* in 1998. The singular circular chromosome is approximately 4.4 Mb and is notably G/C rich (65 %). The *M. tuberculosis* genome contains approximately 4000 protein coding open reading frames (ORFs), of which only 2058 proteins (52 %) have been designated a precise or predicted function (Camus *et al.*, 2002).

The genome sequence identified a large number of genes involved in fatty acid metabolism. Many of which are required for the production of the vast array of lipids, from basic fatty acids to complex mycolic acids which comprise the protective cell wall (Cole *et al.*, 1998). The low permeability of the *M. tuberculosis* cell envelope serves as a barrier and confers intrinsic resistance to some antibiotics (Jarlier and Nikaido, 1990). Genome analysis indicated the presence of a number of drug modifying enzymes, including β -lactamases (Flores *et al.*, 2005) and aminoglycoside acetyltransferases (Ainsa

et al., 1997), as well as putative drug efflux systems, such as ATP-binding cassette (ABC) transporters (Choudhuri *et al.*, 2002, Pasca *et al.*, 2004) and members of the major facilitator superfamily (Takiff *et al.*, 1996, Ainsa *et al.*, 1998, Silva *et al.*, 2001).

M. tuberculosis encodes a number of genes involved in virulence and immunogenicity. Genome analysis identified 170 proteins encoded by two large multi-gene families of unknown function, proposed to be involved in antigenic variation (Cole *et al.*, 1998). The PE and PPE proteins, recognised by N-terminal proline-glutamic acid or proline-proline-glutamic acid motifs, are expressed on the cell surface and recognised as antigens by the host (Delogu and Brennan, 2001, Skeiky *et al.*, 2000, Brennan *et al.*, 2001, Banu *et al.*, 2002, Okkels *et al.*, 2003). Whole-genome comparison of *M. tuberculosis* H37Rv and clinical isolate *M. tuberculosis* CDC1551 revealed single-base substitutions of the PE and PPE families were significantly greater than the genome as a whole (Fleischmann *et al.*, 2002). Likewise, PE and PPE polymorphism between closely related *M. tuberculosis* H37Rv and *M. bovis* is almost 60 % at the protein level (Garnier *et al.*, 2003). Differential expression of PE and PPE genes suggests the ability to alter the antigenic profile during progression of the infection, and may confer tissue or host specificity (Garnier *et al.*, 2003, Voskuil *et al.*, 2004). Another protein family linked to virulence and pathogenicity is the CFP-10/ESAT-6 family. Like the PE and PPE proteins, the function and mechanism of action of the CFP-10/ESAT-6 proteins are yet to be defined (Section 1.5).

1.3.2 Comparative and Functional Genomics

M. tuberculosis is a very successful pathogen, yet the basis of virulence and pathogenesis remains poorly understood. Comparative genomic approaches have been employed to identify genetic differences between members of the *M. tuberculosis* complex, to aid

understanding of virulence and attenuation, and identify potential drug targets and vaccine candidates.

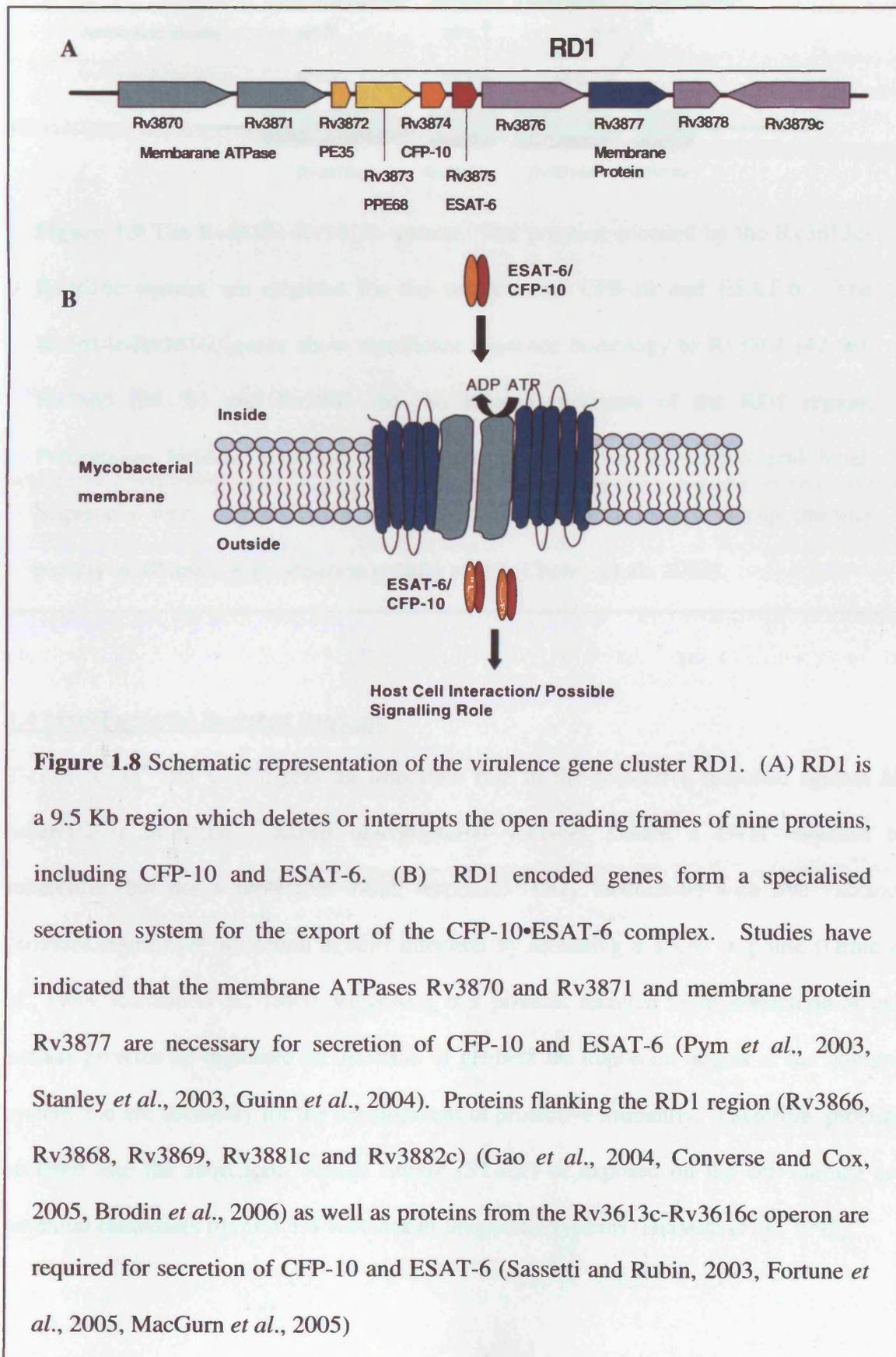
Comparison of *M. tuberculosis*, *M. bovis* and *M. bovis* BCG genomes using subtractive genomic hybridisation and whole genome DNA microarray identified 14 regions deleted from attenuated BCG strains (Mahairas *et al.*, 1996, Behr *et al.*, 1999). Of these regions, only RD1 (Rv3871-Rv3879c) was present in all virulent laboratory and clinical isolates and absent from all BCG daughter strains, clearly indicating a role in TB pathogenesis. Similarly, Harboe *et al.* (1996) and Philipp *et al.* (1996) demonstrated the presence of the RD1 encoded ESAT-6 gene in virulent *M. tuberculosis* and *M. bovis* strains and its absence in *M. bovis* BCG. Comparative analysis between *M. tuberculosis* and *M. microti* identified RD1^{MIC} (Rv3864-Rv3876) which partially overlaps the RD1 deletion observed for *M. bovis* BCG strains (Brodin *et al.*, 2002, Frota *et al.*, 2004).

Functional studies have confirmed the role of RD1 encoded proteins in virulence and pathogenesis. Knock-out studies have shown that deletion of RD1 from *M. tuberculosis* results in decreased virulence. Short-term studies (21 weeks) in mice report that *M. tuberculosis* H37Rv:ΔRD1 shows levels of attenuation similar to BCG (Lewis *et al.*, 2003). Likewise, longer-term studies (> 1 year) with the recombinant H37Rv:ΔRD1 strain also showed reduced virulence, however the knock-out strain was more aggressive than BCG and induced symptoms of active TB (Sherman *et al.*, 2004). Deletion of ESAT-6 from *M. bovis* also reduces virulence, as demonstrated by Wards *et al.* (2000). In comparison, re-introduction of RD1 into *M. bovis* BCG and *M. microti* significantly increases virulence of the recombinant strain, however RD1 knock-in strains do not attain the same level of virulence as *M. tuberculosis* or *M. bovis* (Pym *et al.*, 2002, Majlessi *et al.*, 2005).

The RD1 region contains the genes encoding nine proteins designated Rv3871 to Rv3879c, and includes the immunodominant antigens CFP-10 (Rv3874 or *esxB*) and ESAT-6 (Rv3875 or *esxA*), as well as PE35 (Rv3872) and PPE68 (Rv3873) (Cockle *et al.*, 2002, Mustafa *et al.*, 2002) (Figure 1.8). CFP-10 and ESAT-6 are exported from the mycobacterial cell by a dedicated transport system. The proteins encoded by RD1 and flanking genes form a membrane complex required for the ATP dependant export of CFP-10 and ESAT-6. Rv3870 and Rv3871 (membrane-bound ATPase), and Rv3877 (integral membrane protein) are essential for the secretion of CFP-10 and ESAT-6 (Tekaiia *et al.*, 1999, Hsu *et al.*, 2003, Pym *et al.*, 2003, Stanley *et al.*, 2003, Guinn *et al.*, 2004). A recent study has also shown that deletion of Rv3872 (PE35) results in no expression of CFP-10 and ESAT-6, suggesting that Rv3872 may be involved in the regulation of CFP-10/ESAT-6 expression (Brodin *et al.*, 2006). In contrast, truncation or inactivation of Rv3873 (PPE68) resulted in a slight increase in exported ESAT-6, indicating that this protein could possibly be involved in regulation of secretion of RD1 effectors CFP-10 and ESAT-6 (Brodin *et al.*, 2006). RD1 proteins Rv3876 (proline and alanine rich protein), Rv3878 (alanine rich protein) and Rv3879c (proline rich protein) are not involved in the secretion process (Inwald *et al.*, 2003, Pym *et al.*, 2003, Demangel *et al.*, 2004, Brodin *et al.*, 2006).

Recent studies have also identified a RD1 independent region, containing the genes Rv3613c to Rv3616c, which plays a role in virulence and is necessary for CFP-10/ESAT-6 secretion (Sasseti and Rubin, 2003, Fortune *et al.*, 2005, MacGurn *et al.*, 2005). Rv3613, Rv3614c, Rv3615c and Rv3616c are organised as an operon (Rickman *et al.*, 2005), and show homology to proteins upstream of the RD1 region (Figure 1.9) (MacGurn *et al.*, 2005). Additional studies carried out in *M. smegmatis* and *M. marinum* have indicated that genes flanking the RD1 region (Rv3866, Rv3868, Rv3869, Rv3881c and Rv3882c) may

also be required for secretion of CFP-10 and ESAT-6 (Gao *et al.*, 2004, Converse and Cox, 2005, Brodin *et al.*, 2006).



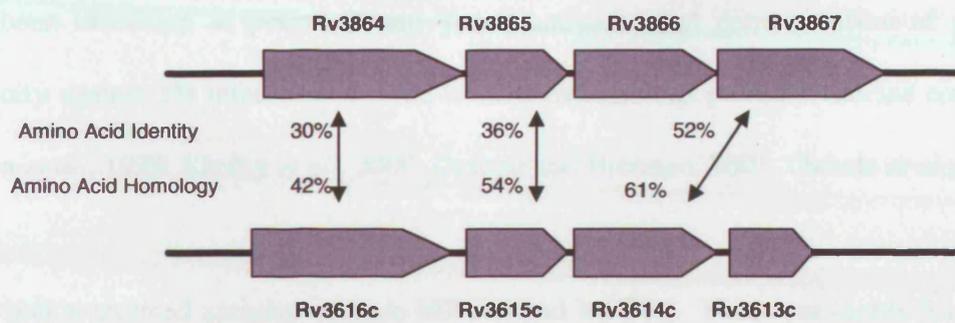


Figure 1.9 The Rv3613c-Rv3616c operon. The proteins encoded by the Rv3613c-Rv3616c operon are required for the secretion of CFP-10 and ESAT-6. The Rv3614c-Rv3616c genes show significant sequence homology to Rv3864 (42 %), Rv3865 (54 %) and Rv3867 (61 %) located upstream of the RD1 region. Percentages indicate sequence identity and homology at the amino acid level. Sequences were aligned using ClustalW with a Gonnet matrix, a gap opening penalty of 10 and a gap extension penalty of 0.2 (Chenna et al., 2003).

1.4 Mycobacterial Secreted Proteins

T-cells (CD4⁺ and CD8⁺) play an important role in the protective response against *M. tuberculosis* infection. Killed mycobacterial vaccines induce a DTH response to tuberculin, but not a protective T-cell response. Only inoculation with live vaccines provides significant protection against infection by mounting a T-cell response (Orme *et al.*, 1988, Romain *et al.*, 1993), suggesting that proteins secreted by mycobacteria or cell surface proteins up-regulated on infection of the host are important targets of the immune system and are necessary for the development of protective immunity. Therefore, proteins secreted into the short term culture filtrate (ST-CF) or exposed on the cell surface are potential candidates for new TB vaccines or diagnostic systems (Horwitz *et al.*, 1995).

Members of the large PE and PPE families appear to be expressed at the cell surface and have been identified as potent B and T-cell antigens, and demonstration of protective immunity against TB infection have led to their inclusion as possible vaccine components (Dillon *et al.*, 1999, Skeiky *et al.*, 2000, Delogu and Brennan, 2001, Okkels *et al.*, 2003).

Other major secreted antigens include MPT83 and MPT70. These are highly homologous proteins produced by *M. tuberculosis*, *M. bovis* (MPB83 and MPB70) and *M. bovis* BCG strains (Hewinson *et al.*, 1996). The genes encoding MPT83 (*Rv2873*) and MPT70 (*Rv2875*) are expressed under identical conditions and form part of a larger operon (*Rv2873* to *Rv2878c*) controlled by the sigma factor *sigK* (Charlet *et al.*, 2005). MPT70 and MPT83 share a high level of sequence homology, 80 % over 163 residues (74 % identity), and are thought to adopt a similar structure, as described for MPB70 (Carr *et al.*, 2003). The main differences between the two proteins are that MPT83 is glycosylated and is localised to the cell surface via a lipid group at the N-terminus, where-as non-glycosylated MPT70 is secreted into the ST-CF (Harboe *et al.*, 1998, Michell *et al.*, 2003). Expression of MPT70 and MPT83 varies between different strains. Both proteins are highly expressed by virulent *M. bovis* and some *M. bovis* BCG strains, where-as other *M. bovis* BCG strains, including BCG Pasteur, and *M. tuberculosis* show low expression of MPT70 and MPT83 (Miura *et al.*, 1983, Wiker *et al.*, 1996). Sequence analysis of *sigK* across the *M. tuberculosis* complex showed that the eight low producing BCG strains possess a mutation in the start codon, which significantly reduces translation of *sigK*, resulting in lower expression of MPB70 and MPB83 (Charlet *et al.*, 2005). However, the difference in MPT70/MPB70 and MPT83/MPB83 expression between *M. tuberculosis* and *M. bovis* is not due to the mutation in *sigK*, but is thought to be the result of unregulated *sigK* activity in *M. bovis*, resulting in constitutive expression of MPB70 and MPB83 (Charlet *et al.*, 2005). Both MPT70/MPB70 and MPT83/MPB83 induce a strong immune

response upon challenge with infection, indicating potential use in vaccines (Hewinson *et al.*, 1996).

A number of other secreted *M. tuberculosis* antigens also identified as potential vaccine components include; MPT64 (Haslov *et al.*, 1995, Elhay *et al.*, 1998); heat shock protein GroES (Barnes *et al.*, 1992, Rosenkrands *et al.*, 1999); the Ag85 complex, including antigens 85A (MPT44), 85B (MPT59) and 85C (MPT45) (Wiker and Harboe, 1992); CFP-10 and ESAT-6.

1.5 CFP-10 and ESAT-6

1.5.1 The CFP-10/ESAT-6 Protein Family

The genes encoding CFP-10 (*Rv3874/esxB*) and ESAT-6 (*Rv3875/esxA*) are located adjacent to each other and are co-transcribed under the control of the same promoter, indicating that they are organised as an operon (Bethet *et al.*, 1998). CFP-10 and ESAT-6 are members of a large gene family, including 23 members arranged at 11 loci within the *M. tuberculosis* genome (Figure 1.10). CFP-10/ESAT-6 family members are characterised by their small size (approximately 100 amino acids), their organisation in pairs within the genome and the presence of the central WXG motif (Pallen, 2002). The majority of family members are preceded by PE/PPE proteins and are often part of a larger gene cluster, found five times in the genome of *M. tuberculosis* (Gey van Pittius *et al.*, 2001). These larger gene clusters are similar to that encompassing the RD1 deletion, which is known to encode essential elements of the secretion machinery required for the export of the CFP-10•ESAT-6 complex (Tekaiia *et al.*, 1999, Hsu *et al.*, 2003, Pym *et al.*, 2003, Stanley *et al.*, 2003, Guinn *et al.*, 2004).

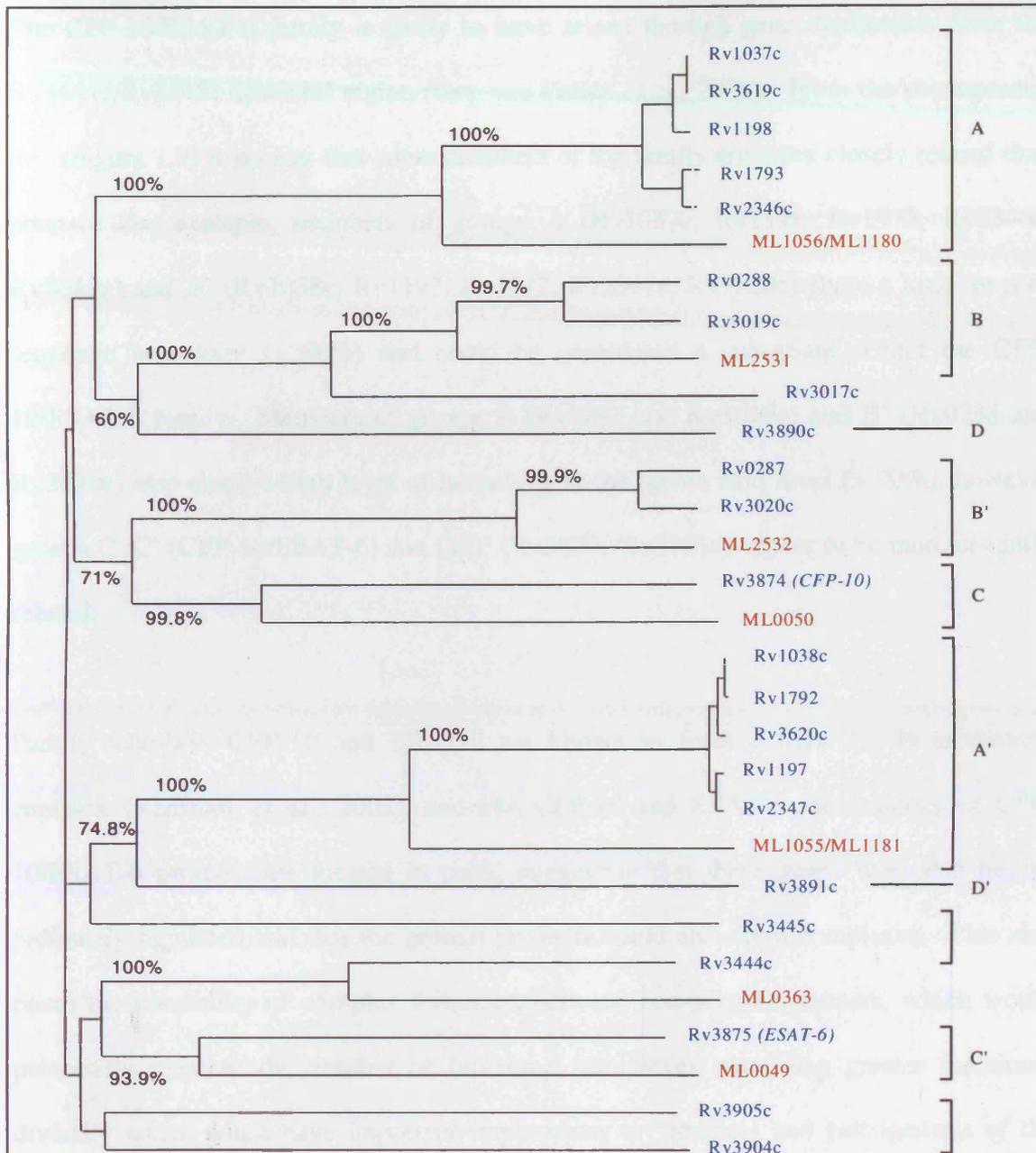


Figure 1.10 Bootstrapped phylogenetic tree for the CFP-10/ESAT-6 family of proteins from *M. tuberculosis* (prefixed by Rv, blue) and their *M. leprae* homologues (prefixed by ML, red) (Prepared by P. Renshaw and reported in Lightbody *et al.*, 2004). There are 23 family members of the CFP-10/ESAT-6 family found in *M. tuberculosis*. Major pairing groups are highlighted by square brackets and labelled (A pairs with A' etc.). The family relationships shown in the tree were generated from optimised protein sequence alignments obtained using the Blosum62 scoring matrix, with a gap opening penalty of 12 and a gap extension penalty of 2 per residue.

The CFP-10/ESAT-6 family is likely to have arisen through gene duplication from the Rv3444c/Rv3445c ancestral region (Gey van Pittius *et al.*, 2001). From the phylogenetic tree (Figure 1.9) it is clear that some members of the family are more closely related than others. For example, members of groups A (Rv1037c, Rv1198, Rv1973, Rv2346c, Rv3619c) and A' (Rv1038c, Rv1197, Rv1972, Rv2347c, Rv3620c) share a high level of sequence homology (> 90%) and could be considered a sub-group within the CFP-10/ESAT-6 family. Members of groups B (Rv0287 and Rv3019c) and B' (Rv0288 and Rv3020c) also share a high level of homology at the amino acid level (> 90%), however groups C /C' (CFP-10/ESAT-6) and D/D' (Rv3890c/Rv3891c) appear to be more distantly related.

Family members CFP-10 and ESAT-6 are known to form a tight 1:1 heterodimeric complex (Renshaw *et al.*, 2002), and like CFP-10 and ESAT-6 the majority of CFP-10/ESAT-6 proteins are located in pairs, suggesting that these genes may also be coordinately regulated and that the protein products could also form complexes. This also raises the possibility of complex formation between non-genome partners, which would potentially increase the number of functional complexes, providing greater functional diversity which could have important implications in virulence and pathogenesis of the mycobacteria. Therefore, understanding the rules governing complex formation would provide an important insight into the CFP-10/ESAT-6 protein family.

CFP-10/ESAT-6 homologues have been identified in eight other species of mycobacteria, including *M. leprae*, as well as closely related *Corynebacterium diphtheriae* and *Streptomyces coelicolor* and the more distantly related Gram positive organism *Staphylococcus aureus* (Gey van Pittius *et al.*, 2001, Burts *et al.*, 2005). The recently published genome of *M. leprae* revealed an extreme case of reductive evolution (Cole *et*

al., 2001). *M. leprae* is thought to contain the minimum gene set required to remain pathogenic, therefore conservation of CFP-10 and ESAT-6 family members indicates their importance to the pathogenic mycobacteria. CFP-10 and ESAT-6 are individually conserved in *M. leprae* (ML0049/ML0050) indicating their significance in virulence and pathogenesis. The *M. leprae* proteins ML2531/ML2352 substitute for both the Rv0287/Rv0288 and Rv3019c/Rv3020c pairs, and an identical pair of *M. leprae* proteins substitute for the five pairs of *M. tuberculosis* proteins in groups A/A' (Figure 1.9), which may indicate some functional redundancy.

1.5.2 Potent T-cell Antigens

The CFP-10/ESAT-6 family includes a number of immunodominant antigens found in the ST-CF of *M. tuberculosis*, including CFP-10/ESAT-6, Rv0287/Rv0288, Rv3019c and members of the A/A' groups (Sorensen *et al.*, 1995, Alderson *et al.*, 2000, Dillon *et al.*, 2000, Rosenkrands *et al.*, 2000, Skjot *et al.*, 2000, Skjot *et al.*, 2002, Mattow *et al.*, 2003). Both CFP-10 and ESAT-6 induce DTH responses in a range of TB infected hosts, including guinea pigs, cattle and humans (Elhay *et al.*, 1998, Colangeli *et al.*, 2000, Van Pinxteren *et al.*, 2000, Brusasca *et al.*, 2001). The immune responses of both humans and cattle to CFP-10 and ESAT-6 are shown to be highly specific for TB infection. T-cell proliferation and IFN γ assays demonstrated that CFP-10 and ESAT-6 are recognised by a large proportion (> 70 %) of TB infected patients or cattle but not by BCG vaccinated or healthy controls (Buddle *et al.*, 1999, Mustafa *et al.*, 1998, Ravn *et al.*, 1999, Arend *et al.*, 2000, Skjot *et al.*, 2000, Mustafa *et al.*, 2002, Vordermeier *et al.*, 2002). Recent research has led to the development of two commercial IFN γ assays which use CFP-10 and ESAT-6; QuantiFERON-TB Gold (Cellestis Limited, Australia), a whole blood assay which uses an enzyme-linked immunosorbent assay (ELISA) to measure IFN γ production and T SPOT-TB (Oxford Immunotec, UK) which uses peripheral blood mononuclear cells

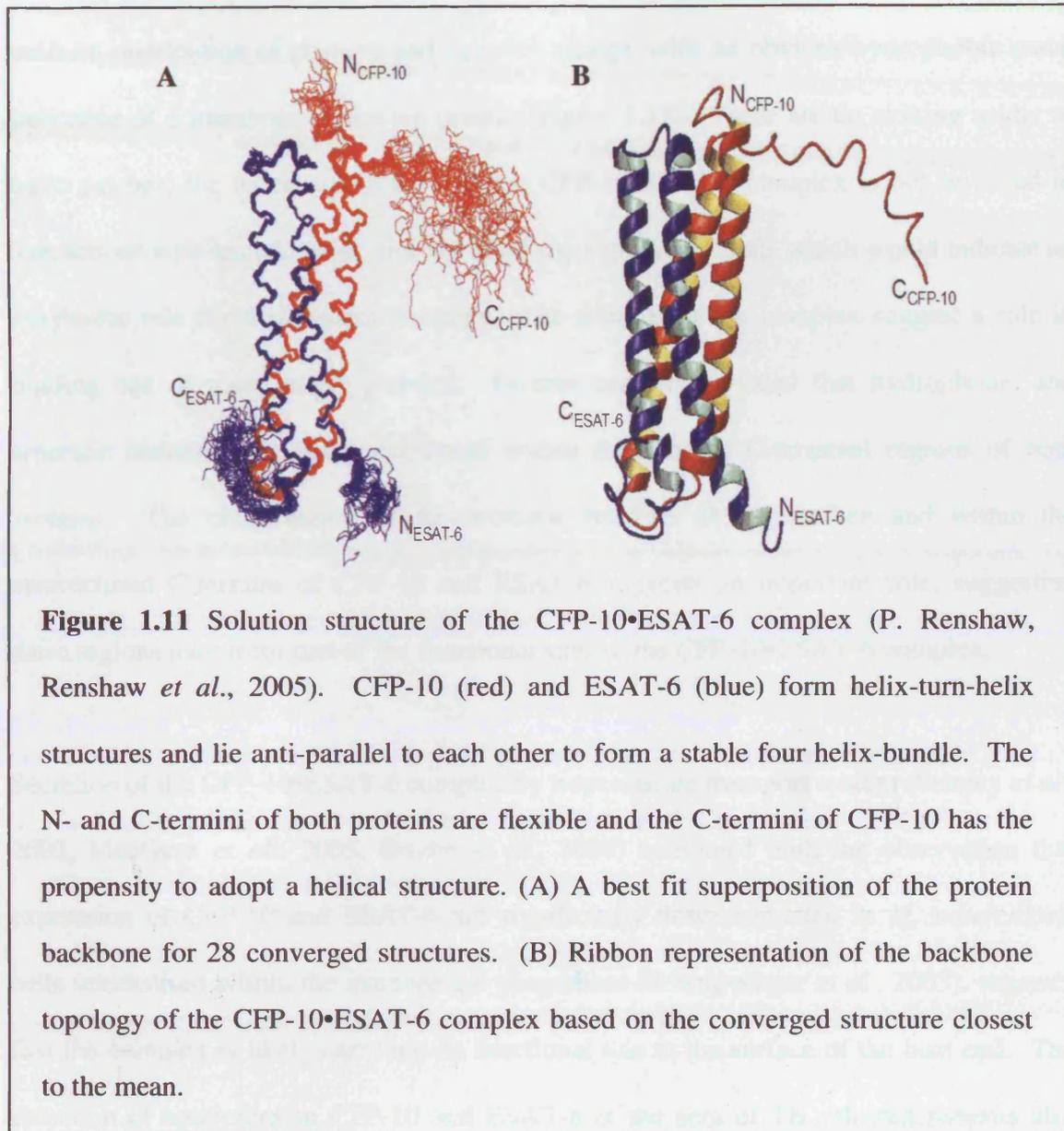
(PBMCs) and detects the number of T-cells producing IFN γ by an enzyme-linked immunospot assay (ELISPOT) (Pai *et al.*, 2005).

Vaccine research has also included CFP-10 and ESAT-6 (Section 1.2.2), however, the use of either of these antigens in a vaccine would prevent their use in diagnostic assays. Therefore vaccine research has investigated the use of other ESAT-6/CFP-10 family proteins, such as Rv0288 (Dietrich *et al.*, 2005) and Rv3019c (Hogarth *et al.*, 2005), as an alternative to ESAT-6 in subunit vaccines. Collins *et al.* (2003) have also investigated the use of attenuated *M. bovis* strains with and without expression of ESAT-6.

1.5.3 Structure and Function of CFP-10 and ESAT-6

Neither CFP-10 nor ESAT-6 shows any significant sequence similarity to proteins of known function. Previous work has shown that CFP-10 and ESAT-6 form a tight 1:1 heterodimer, which is considered to be the functional form of these proteins (Renshaw *et al.*, 2002, Brodin *et al.*, 2005). Okkels *et al.* (2004) reported the presence of various ESAT-6 species in the ST-CF of *M. tuberculosis*, including full length ESAT-6, N-terminal acetylated ESAT-6 and truncated C-terminal ESAT-6. N-terminal acetylation is uncommon in bacterial proteins and the reason for this post-translational modification is unknown. In the study described by Okkels *et al.* (2004) N-terminal acetylation of ESAT-6 (at Thr₂) was reported to abolish the interaction with CFP-10. Based on the recently reported structure of the CFP-10•ESAT-6 complex (Renshaw *et al.*, 2005), the failure of the acetylated ESAT-6 to bind CFP-10 is unlikely to be structural, as the N-terminus of ESAT-6 is not involved in the interface with CFP-10 (Figure 1.11). Further investigation by Brodin *et al.* (2005) demonstrated that ESAT-6 with a T2H mutation is secreted as normal suggesting that the post-translational modification of ESAT-6 plays no obvious functional role. The structure of the complex and intrinsic tryptophan fluorescence studies

described in this thesis agrees with the report by Okkels *et al.* (2004), which indicates that the 11 C-terminal residues of ESAT-6 are not necessary for interaction with CFP-10.



The function of the CFP-10•ESAT-6 complex is yet to be determined, however the recently acquired solution structure of the complex indicates a potential role in protein interactions. The structure shows that both CFP-10 and ESAT-6 adopt a helix-turn-helix hairpin structure in the complex, which are arranged anti-parallel to each other forming a stable four helix bundle (Renshaw *et al.*, 2005) (Figure 1.11). It has previously been

proposed that the CFP-10•ESAT-6 complex possesses cell lysis activity (Hsu *et al.*, 2003), mediated via the formation of pores in the membrane, however, the surface features of the complex argue against this proposal. The surface of the complex demonstrates a fairly uniform distribution of positive and negative charge, with no obvious hydrophobic patch indicative of a membrane spanning protein (Figure 1.12). There are no striking acidic or basic patches, the latter suggesting that the CFP-10•ESAT-6 complex is not involved in interactions with nucleic acids, nor are there any significant clefts which would indicate an enzymatic role for this protein complex. The features of the complex suggest a role in binding one or more target proteins. Further analysis revealed that hydrophobic and aromatic residues are highly conserved within the flexible C-terminal regions of both proteins. The conservation of the aromatic residues at the surface and within the unstructured C-termini of CFP-10 and ESAT-6 suggests an important role, suggesting these regions may form part of the functional sites of the CFP-10•ESAT-6 complex.

Secretion of the CFP-10•ESAT-6 complex by a specialised transport system (Stanley *et al.*, 2003, MacGurn *et al.*, 2005, Brodin *et al.*, 2006) combined with the observation that expression of CFP-10 and ESAT-6 are significantly down-regulated in *M. tuberculosis* cells internalised within the macrophage phagosome (Schnappinger *et al.*, 2003), suggests that the complex is likely carry out its functional role at the surface of the host cell. The detection of antibodies to CFP-10 and ESAT-6 in the sera of TB infected patients also indicates that these antigens are present outside infected macrophages (Brusasca *et al.*, 2001). The fluorescence microscopy assays reported in this thesis strongly suggest that the CFP-10•ESAT-6 complex binds to a specific receptor on the surface of monocyte lineage cells.

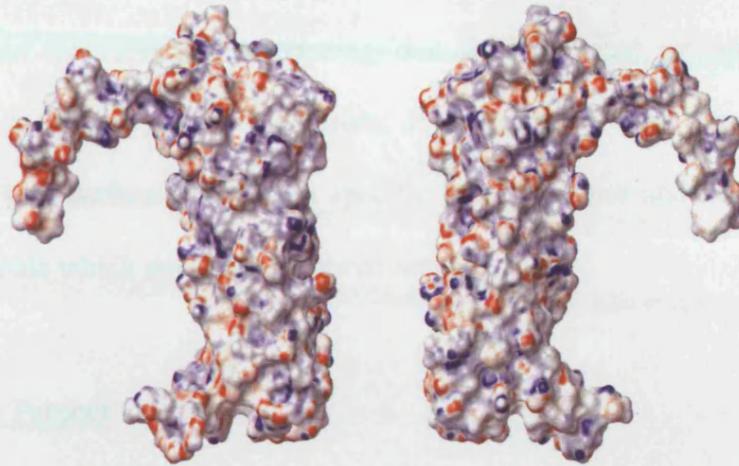


Figure 1.12 Space filled view of the surface features of the CFP-10•ESAT-6 complex. Both faces of the complex are shown, with the right hand view in the same orientation as figure 1.10. The surface is coloured according to electrostatic potential (red - negative charge, blue - positive charge, neutral - white). The surface features of the complex demonstrate a fairly uniform distribution of positive and negative charge, with no obvious hydrophobic patches indicative of a membrane spanning protein. There are no striking acidic or basic patches, nor are there any significant clefts which would indicate an enzymatic role. The features of the complex suggest a role in binding one or more target proteins.

However, despite the current interest in the CFP-10•ESAT-6 complex, the molecular function and mechanism of action of this protein complex remain unknown. A recent study by Volkman *et al.*, (2004) suggested that RD1 secreted effectors, CFP-10 and ESAT-6, mediate aggregation of host macrophages, facilitating the intracellular spread of mycobacteria. It is possible that the RD1 effectors interact with components of host cell signalling pathways to induce macrophage aggregation. Previous studies with individual CFP-10 and ESAT-6 proteins have also suggested a role in modulation of host cell activity, possibly involving cytokine production (Trajkovic *et al.*, 2002, Singh *et al.*, 2005). More

recently Brodin *et al.* (2006) reported the increased uptake of mycobacteria expressing the RD1 region into macrophages, suggesting that RD1 secreted proteins may play an important role in phagocytosis. Therefore, it is suggested that the CFP-10•ESAT-6 complex binds to a surface receptor on specific host cell-types and is likely to act as a signalling molecule which modulates host cell activity.

1.6 Aims of the Project

The work described in this thesis forms part of a wider programme which aims to structurally and functionally characterise proteins implicated in the virulence and pathogenesis of *M. tuberculosis* and *M. bovis*. These include the CFP-10/ESAT-6 family, MPT70 (MPB70) and members of the Rv3613c-Rv3616c operon. The primary aim of this project was to investigate the protein-protein interactions of the *M. tuberculosis* complex CFP-10/ESAT-6 protein family, where the main objectives were to (i) characterise complex formation between members of the CFP-10/ESAT-6 family to and (ii) to investigate the binding of the CFP-10•ESAT-6 complex to the surface of host cells.

Yeast two-hybrid studies were employed to investigate interactions between genome pairs of the CFP-10/ESAT-6 protein family and, more importantly to identify which members of the family could cross-talk with non-genome partners, as interactions between non-genome partners could greatly increase the number of complexes and potentially enhance the functional diversity of this protein family. These studies along with helical wheel representations and multiple sequence alignments were used to establish the rules which are likely to govern complex formation within the CFP-10/ESAT-6 protein family.

Fluorescence microscopy assays were used to determine the binding of Alexa Fluor 546 (Molecular Probes) labelled CFP-10•ESAT-6 complex to the surface of host cells,

including monocyte and fibroblast-type cells. To eliminate non-specific interactions and to confirm specificity, cells were exposed to both labelled and unlabelled complex. Finally, based on the structural studies the flexible C-termini of both CFP-10 and ESAT-6 were removed to establish whether these regions play a functional role in binding to the surface of host cells. Complex formation between full length and C-terminal truncated proteins was analysed by intrinsic tryptophan fluorescence before investigating binding of CFP-10•ESAT-6 truncated complexes to the surface of monocytic lineage host cells.

Chapter 2

Specificity of Interactions between Members of the CFP-10/ESAT-6 Family

2.1 Introduction

CFP-10 and ESAT-6 are members of a large protein family within the *M. tuberculosis* complex, comprising 23 members usually found in pairs within the genome (Figure 2.1). Previous work by Renshaw *et al.* (2002) showed that CFP-10 and ESAT-6 form a tight 1:1 heterodimeric complex and suggested that other members of the family may also be able to form complexes with their genome partners. This also raises the possibility of complex formation between non-genome partners, which would provide greater functional flexibility for CFP-10/ESAT-6 protein family. A yeast two-hybrid approach was used to investigate complex formation between genome and non-genome partners within the CFP-10/ESAT-6 family.

The yeast two-hybrid system is based on the understanding that most eukaryotic transcription factors are composed of two distinct functional domains, a DNA binding domain (DBD), which targets the transcription factor to a specific promoter, and the transcription activation domain (TAD) which is involved in the assembly of the transcription complex (Hope and Struhl, 1986, Ma and Ptashne 1988). The DBD and TAD do not need to be physically linked, however neither domain can activate transcription alone, therefore the DBD and TAD must co-localise at the promoter to allow transcription to occur.

The yeast GAL4 protein (881 amino acids) activates transcription of genes involved in galactose catabolism (*GAL1*, *GAL2*, *GAL7*, *GAL10* and *MEL1*) (Johnston, 1987), and is

negatively regulated by the gene product of *GAL80* which binds to the C-terminal 30 residues of GAL4 inhibiting gene activation (Johnston *et al.*, 1987, Ma and Ptashne, 1987a).

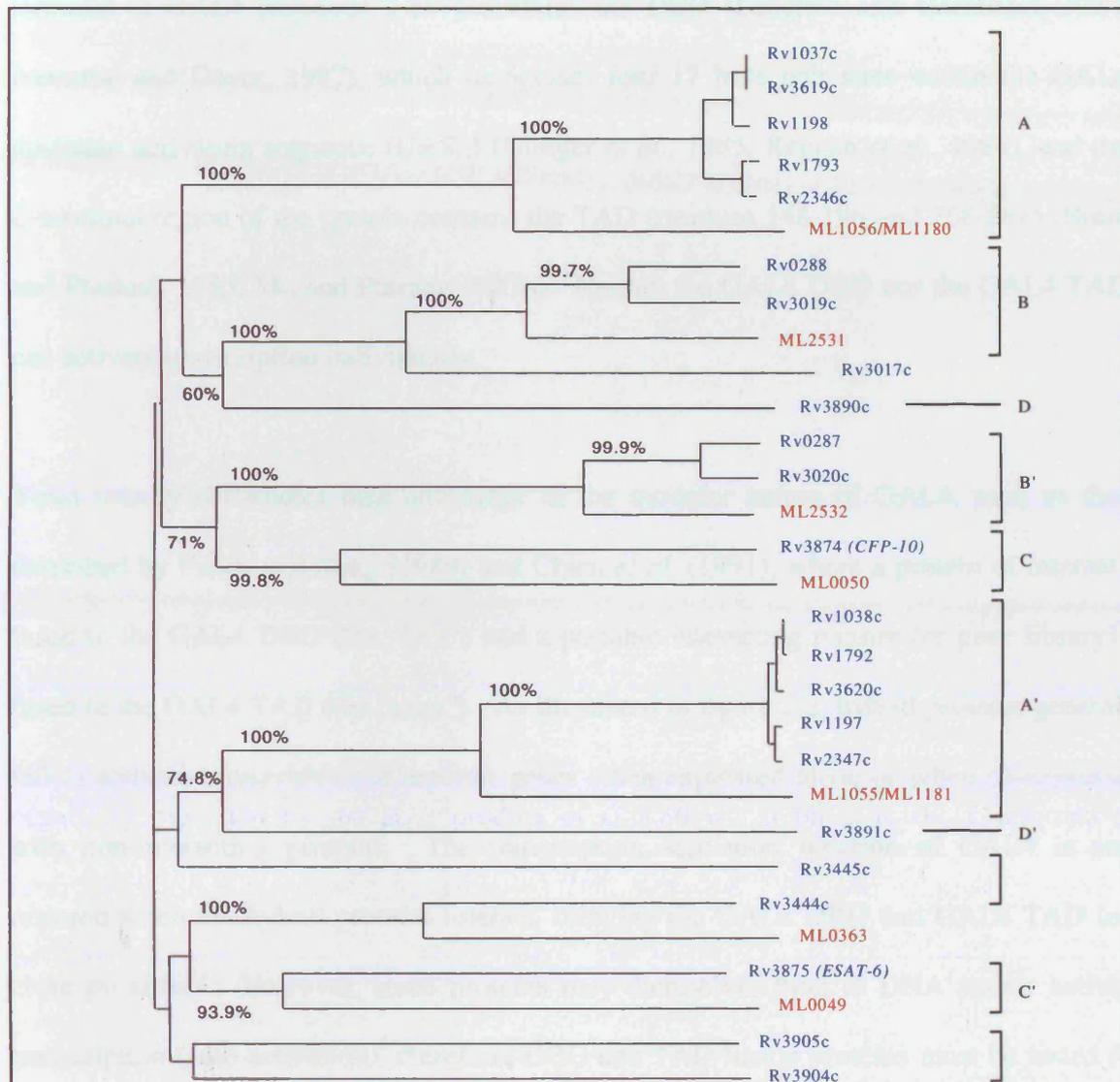


Figure 2.1 Phylogenetic tree for CFP-10/ESAT-6 family of proteins from *M. tuberculosis* (prefixed by Rv, blue) and their *M. leprae* homologues (prefixed by ML, red) (Prepared by P. Renshaw, reported in Lightbody *et al.*, 2004). There are 23 family members of the CFP-10/ESAT-6 family found in *M. tuberculosis*. Major pairing groups are highlighted by square brackets and labelled (A pairs with A' etc.). Bootstrap values (%) are indicated for the major branch points in the tree. To generate the family relationships shown in the tree, the protein sequences were aligned using the Blosum62 scoring matrix.

Therefore, yeast host strains used in GAL4-based systems must carry deletions of the *GAL4* and *GAL80* genes to prevent interference by native GAL4 and GAL80. The N-terminal of GAL4 (residues 1-147) contains the DBD (Laughon and Gesteland, 1984, Johnston and Dover, 1987), which recognises four 17 base pair sites within the GAL4 upstream activating sequence (UAS_G) (Giniger *et al.*, 1985, Keegan *et al.*, 1986), and the C-terminal region of the protein contains the TAD (residues 148-196 and 768-881) (Brent and Ptashne, 1985, Ma and Ptashne 1987b). Neither the GAL4 DBD nor the GAL4 TAD can activate transcription individually.

Yeast two-hybrid studies take advantage of the modular nature of GAL4, such as those described by Fields and Song (1989) and Chien *et al.* (1991), where a protein of interest is fused to the GAL4 DBD (the “bait”) and a possible interacting partner (or gene library) is fused to the GAL4 TAD (the “prey”). As illustrated in figure 2.2, hybrid proteins generally fail to activate transcription of reporter genes when expressed alone or when co-expressed with non-interacting proteins. The transcription activation function of GAL4 is only restored when the hybrid proteins interact, bringing the GAL4 DBD and GAL4 TAD into close proximity. However, some proteins may themselves bind to DNA and/or activate transcription (auto-activation), therefore, DBD and TAD fusion proteins must be tested for auto-activation of reporter genes.

Most yeast host strains used for yeast two-hybrid studies carry the *lacZ* reporter gene encoding the enzyme β -galactosidase. Expression of *lacZ* can be detected and quantified by colorimetric β -galactosidase assays (Serebriiskii and Golemis, 2000, Rupp, 2002). Solid support assays use X-gal (5-bromo-4-chloro-3-indolyl- β -D-galactoside) as the substrate. X-gal is cleaved by β -galactosidase which produces a blue precipitate, allowing positive interactions to be easily identified by the development of blue colonies.

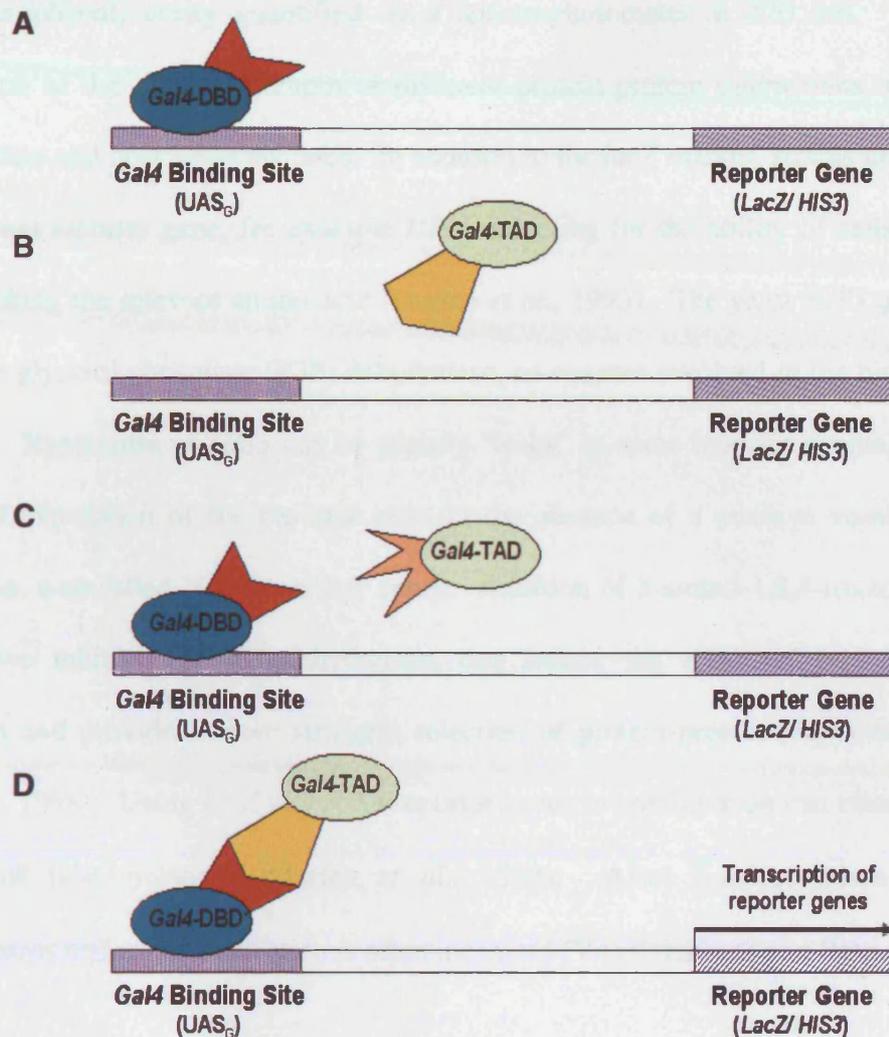


Figure 2.2 The yeast two-hybrid system. (A) A protein of interest fused to the *GAL4* DNA Binding Domain (DBD) binds to DNA but is unable to activate transcription of the reporter genes. (B) A *GAL4* Transcription Activation Domain (TAD) fusion protein contains the transcription activation region but cannot initiate transcription because it fails to bind to the *GAL4* binding site (UAS_G). (C) No interaction between the hybrid proteins means that there is no co-localisation of the *GAL4* DBD and TAD, and subsequently no transcription of the reporter genes. (D) A functional *GAL4* activator is created when the fusion proteins interact bringing the *GAL4* DBD and TAD into close proximity, leading to transcription of the *lacZ* and *HIS3* reporter genes.

β -galactosidase activity can also be quantified by liquid assays, where hydrolysis of the substrate ONPG (*ortho*-nitrophenyl- β -D-galactopyranoside) yields a yellow product ONP

(*ortho*-nitrophenol) easily quantified by a spectrophotometer at 420 nm. This allows comparison of the relative strength of different protein-protein interactions, provided the same vectors and host strain are used. In addition to the *lacZ* marker, strains may also carry a nutritional reporter gene, for example *HIS3*, selecting for the ability of cells to grow on media lacking the relevant amino acid (Durfee *et al.*, 1993). The yeast *HIS3* gene encodes imidazole glycerol phosphate (IGP) dehydratase, an enzyme involved in the biosynthesis of histidine. Expression of *HIS3* can be slightly 'leaky' in some reporter strains, resulting in low level expression of the reporter gene in the absence of a positive yeast two-hybrid interaction, a so called 'false positive' result. Addition of 3-amino-1,2,4-triazole (3-AT), a competitive inhibitor of IGP-dehydratase, can reduce the effect of background *HIS3* activation and provide a more stringent selection of protein-protein interactions (Kishore and Shah, 1988). Using *lacZ* and *HIS3* reporter genes in combination can eliminate a large number of false positives (Durfee *et al.*, 1993). Also, β -galactosidase activity of transformants under *HIS3* selection is often increased (Van Criekinge and Beyaert, 1999).

Preliminary yeast two-hybrid experiments were conducted with CFP-10 and ESAT-6 to assess whether this system could detect the interaction between CFP-10 and ESAT-6 as previously described by Renshaw *et al.* (2002). Subsequent studies were used to further characterise complex formation within the CFP-10/ESAT-6 protein family. Family members Rv0287/Rv0288 and Rv3019c/Rv3020c were selected, along with CFP-10/ESAT-6, to investigate pair wise interactions and complex formation between non-genome partners. The closely related pairs Rv0287/Rv0288 and Rv3019c/Rv3020c share over 95 % amino acid sequence homology, and are located within groups B and B' in the CFP-10/ESAT-6 phylogenetic tree (Figure 2.1), making these pairs likely candidates for complex formation between non-genome partners. CFP-10 and ESAT-6 are distantly related (groups C and C'), suggesting that non-genome partner interactions are less likely

to occur. Rv0287, Rv0288 and Rv3019c, like CFP-10 and ESAT-6 are immunodominant antigens secreted by *M. tuberculosis* and have been identified as possible TB vaccine components (Rosenkrands *et al.*, 2000), therefore defining the rules governing complex formation within this family may lead to a better understanding of the functional roles of CFP-10/ESAT-6 family members.

2.2 Materials and Methods

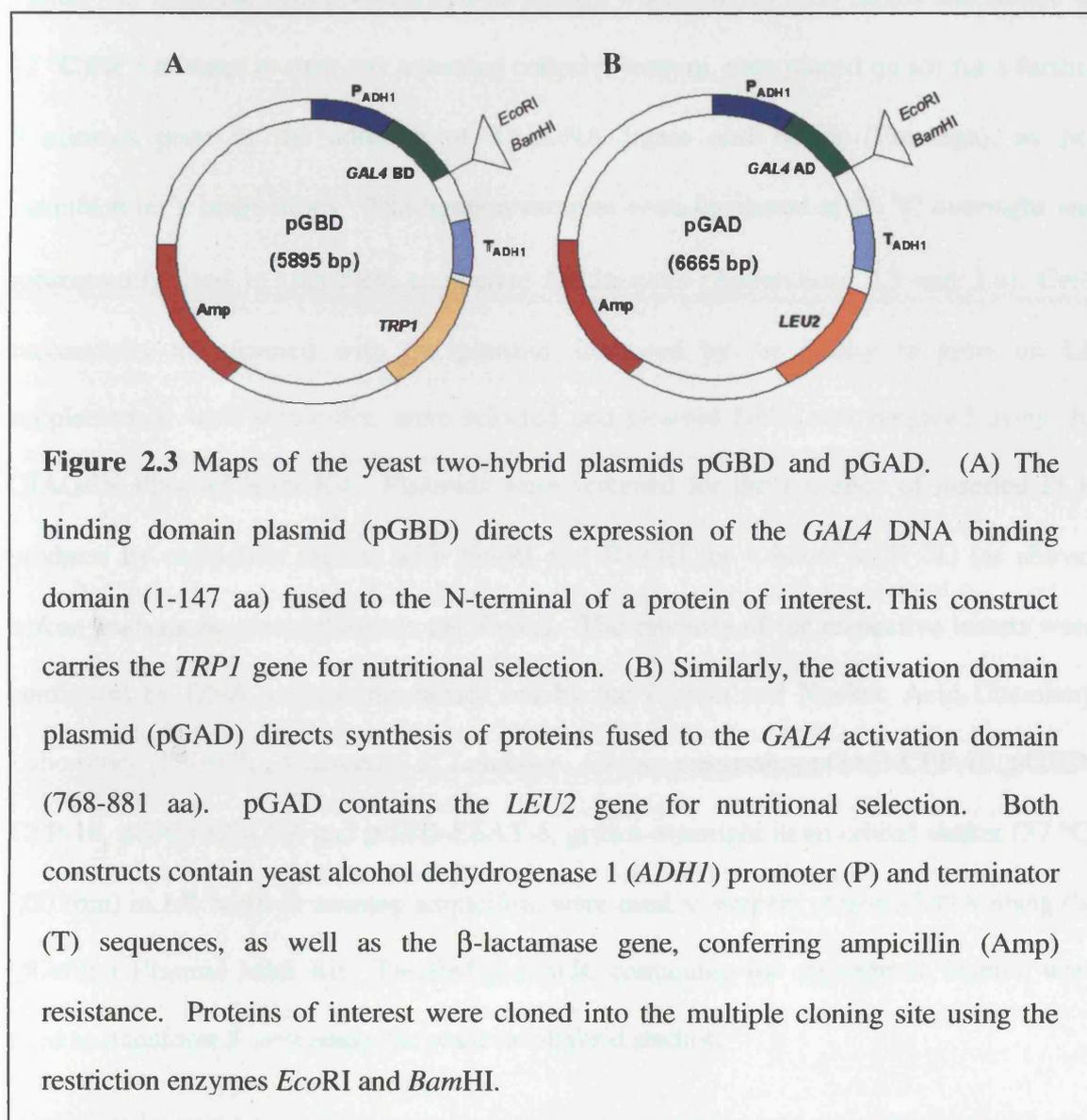
2.2.1 Strains, Plasmids and Media

Escherichia coli strain DH5 α (Appendix 1.1), was used for the transformation, ligation and amplification of plasmids. Cells were grown at 37 °C in Luria-Bertani (LB) medium supplemented with 100 μ g/ml ampicillin (Appendix 1.2). The yeast two-hybrid screens were carried out using *Saccharomyces cerevisiae* strain CG1945 (Appendix 1.1), a derivative of the HF7c as described by Feilotter *et al.* (1994). This strain contains both the *lacZ* and *HIS3* reporter genes, as described in section 2.1, and is auxotrophic for both tryptophan and leucine. The DNA binding domain vector (pGBD) carries the *TRP1* gene and the activation domain vector (pGAD) the *LEU2* gene (Figure 2.3) (Bartel *et al.*, 1993, James *et al.*, 1996), allowing nutritional selection of transformed colonies carrying both vectors. *S. cerevisiae* were grown at 30 °C on either YPD (yeast extract/peptone/dextrose) media or on synthetic dropout (SD) media as required (Appendix 1.3).

2.2.2 Construction of Yeast Two-Hybrid Vectors

Plasmids pGAD and pGBD were supplied by collaborators at The National Institute for Medical Research (NIMR), Mill Hill. PCR was used to amplify the CFP-10 (EsxB) and ESAT-6 (EsxA) coding regions from the bacterial artificial chromosome (BAC) Rv414 (Brosch *et al.*, 1998) template using *Pfu* polymerase (Promega). Forward primers were designed with an *EcoRI* restriction site and reverse primers with a *BamHI* site (Appendix

2). PCR reactions were carried out in 50 μ l volumes (Promega protocol) using a Biometra TRIO-thermablock programmed with the following protocol; initial denaturation at 94 $^{\circ}$ C for 1 minute 30 seconds, followed by 30 cycles of denaturation at 94 $^{\circ}$ C for 30 seconds, annealing at 55 $^{\circ}$ C for 1 minute and extension at 72 $^{\circ}$ C for 1 minute, with a final extension time of 5 minutes at 72 $^{\circ}$ C. PCR products were analysed by electrophoresis, as described in appendix 3.1. PCR products were purified using the PCR Clean-up Kit (Qiagen).



The yeast two hybrid vectors, pGAD and pGBD, and the purified PCR products, CFP-10 and ESAT-6, were restricted with *EcoRI* and *BamHI* (Promega) for 4 hours at 37 $^{\circ}$ C (Appendix 3.2). Both the vectors and the PCR products were purified using the QIAquick

Gel Extraction Kit (Qiagen). The concentration of the recovered vectors were analysed by ultraviolet (UV) absorption at 260 nm using a Cary Bio 300 UV-visible spectrophotometer (Varian). The concentrations of the PCR products were estimated by comparison to the intensities of bands from the 100 bp DNA Marker (Promega).

Ligation reactions were carried out in 20 µl volumes with a 3:1 molar ratio of insert to vector. Initially, the insert, vector and de-ionised water (dH₂O) were mixed and heated to 42 °C for 5 minutes to melt any annealed cohesive termini, then placed on ice for a further 5 minutes prior to the addition of T4 DNA ligase and buffer (Promega), as per manufacturer's instructions. The ligation samples were incubated at 16 °C overnight and subsequently used to transform competent DH5α cells (Appendices 3.3 and 3.4). Cells successfully transformed with the plasmid, indicated by the ability to grow on LB supplemented with ampicillin, were selected and plasmid DNA was prepared using the QIAGEN Plasmid Mini Kit. Plasmids were screened for the presence of inserted PCR products by restriction digests with *EcoRI* and *BamHI* for 4 hours at 37 °C (as above) before analysis by electrophoresis (as above). The integrity of the respective inserts were confirmed by DNA sequencing carried out by the Protein and Nucleic Acid Chemistry Laboratory (PNACL), University of Leicester. Clones possessing pGAD-CFP-10, pGBD-CFP-10, pGAD-ESAT-6 and pGBD-ESAT-6, grown overnight in an orbital shaker (37 °C/ 200 rpm) in LB broth containing ampicillin, were used to prepare plasmid DNA using the QIAfilter Plasmid Midi Kit. Purified plasmids, containing the appropriate inserts, were used to transform *S. cerevisiae* for yeast two-hybrid studies.

Identical procedures were employed to produce yeast two-hybrid constructs containing the genes *Rv0287* (*esxG*), *Rv0288* (*esxH*) and *Rv3020c* (*esxS*). All primers used are listed in Appendix 2. Amplification of *Rv3019c* (*esxR*) was performed using the Advantage GC 2

PCR Kit (BD Biosciences). PCR reactions were set up in 50 µl volumes following the manufacturer's instructions. The PCR protocol included an initial denaturation step of 94 °C for 3 minutes, followed by 30 cycles of 94 °C for 30 seconds, 55 °C for 1 minute, 72 °C for 1 minute and a final extension of 5 minutes at 72 °C. Rv3019c was cloned into pGBD and pGAD as described previously for CFP-10 and ESAT-6.

2.2.3 Transformation of *S. cerevisiae*

S. cerevisiae were transformed using the lithium acetate/single-stranded carrier DNA/polyethylene glycol (LiAc/ss-DNA/PEG) method (Gietz and Woods, 2002). Fresh yeast colonies were used to inoculate 30 ml YPD broth (Appendix 1.3), which was incubated overnight at 30 °C in an orbital shaker at 200 rpm. 15 ml from the overnight culture was inoculated into 35 ml of fresh YPD and incubated for 5 hours at 30 °C in an orbital shaker (200 rpm). The cells were pelleted by centrifugation at 3000 rpm for 15 minutes at 4 °C then washed in 25 ml sterile dH₂O. The centrifugation step was repeated and the supernatant discarded. The cells were resuspended in 1 ml 100 mM LiAc before centrifugation at 13,000 rpm for 1 minute at room temperature. The cell pellet was resuspended in 400 µl 100 mM LiAc, vortexed then aliquoted into 50 µl samples. The cells were pelleted by centrifugation (13,000 rpm for 30 seconds at room temperature) and the LiAc supernatant removed. Meanwhile, single stranded herring sperm DNA (Sigma) (5 mg/ml) in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) was incubated at 100 °C for 5 minutes then chilled on ice prior to use. The transformation mix, composed of 240 µl 50 % (w/v) PEG, 36 µl 1M LiAc, 20 µl single stranded DNA, 1 µg of each plasmid and dH₂O to a final volume of 360 µl was added to the pelleted yeast cells and vortexed until fully mixed. The transformations were incubated at 30 °C for 30 minutes followed by heat shock at 42 °C for 30 minutes. The cells were then pelleted by centrifugation (13,000 rpm for 30 seconds at room temperature) and the transformation mix removed. Transformed

yeast cells were resuspended in 200 μ l sterile dH₂O and the cell suspension plated out on the appropriate SD dropout media (Appendix 1.3). The plates were incubated at 30 °C until colonies were clearly visible (3-10 days).

2.2.4 β -galactosidase Assays

2.2.4.1 Colony Filter Lift Assays

LacZ reporter gene activity was initially detected by X-gal filter lift assays as described by Breeden and Naysmith (1985). The transformed yeast cells were grown on appropriate SD plates (as above) until growth was clearly visible. The colonies were transferred to a nitrocellulose membrane (Stratagene) and permeabilised by submersion in liquid nitrogen. The frozen filters were allowed to thaw before repeating the action. Thawed membranes were placed colony side up on Whatman filter paper pre-soaked with Z-buffer/X-gal/ β -mercaptoethanol solution (Appendix 1.4). The assays were incubated at room temperature and the development of blue colonies was recorded. Reactions were stopped by transferring the membrane onto a Whatman filter soaked with 1 M Na₂CO₃ for at least 10 minutes. Membranes were transferred to a fresh Petri dish and allowed to dry.

2.2.4.2 Quantitative Liquid Assays

Specific β -galactosidase activity was quantified by liquid assays of cell free extracts, using ONPG as the substrate (Le Douarin *et al.*, 2001). Successfully transformed yeast colonies were grown in 15 ml SD media lacking selected amino acids overnight at in an orbital shaker at 30 °C and 200 rpm. The overnight cultures were centrifuged at 3000 g for 5 minutes at 4 °C. The supernatant was discarded and the cell pellets resuspended in 1 ml sterile dH₂O. The cells were pelleted at 13,000 rpm for 30 seconds at room temperature, the supernatant discarded and the cell pellet resuspended in 150 μ l Z-buffer (Appendix 1.4). Acid washed glass beads (212-300 μ m and 425-600 μ m) (Sigma) were added to the

cell suspension until the level reached just below the meniscus of the liquid. The samples were placed on ice and vortexed vigorously for 30 seconds before being returned to the ice. This was repeated four times for each sample. Samples were centrifuged at 13,000 rpm for 15 minutes at room temperature and the cell free protein extract was transferred to a fresh tube. 20 μ l of the protein extract was transferred to a fresh tube and adjusted to 500 μ l with the addition of Z-buffer. These samples were vortexed and equilibrated at 30 °C for 5 minutes. Reactions were initiated by the addition of 100 μ l of 4 mg/ml ONPG in Z-buffer. The samples were mixed and the reaction was timed until a pale yellow colour had developed, then the reaction was stopped by the addition of 250 μ l 1 M Na₂CO₃. Control reactions and negative reactions were stopped after 30 minutes by the addition of 250 μ l 1 M Na₂CO₃. The optical density of the samples was recorded at 420 nm. The protein concentration of the cell free lysate was measured using Bradford reagent (Bio-Rad), as per the manufacturer's instructions. The specific activity of the protein extract was calculated using the following formula (Le Douarin *et al.*, 2001):

$$\text{Specific Activity (nmol/mg/min)} = \frac{\text{OD}_{420} \times 0.85}{0.0045 \times [\text{Protein}] \times \text{Volume} \times \text{Time}}$$

OD₄₂₀ is the optical density of the product *o*-nitrophenol at 420 nm and the factor 0.85 corrects for the reaction volume of the ONPG assay. 0.0045 is the optical density of a 1 nmol/ml solution of *o*-nitrophenol (molar extinction coefficient is 4500 M⁻¹ cm⁻¹ at 420 nm), [Protein] is the protein concentration of the cell free extract in mg/ml, Volume is the volume of protein extract used in the ONPG assay in ml and Time is the time taken in minutes before the ONPG reaction was stopped.

Statistical analysis was performed using Sigma Plot 8.0 (SPSS). Differences in β -galactosidase activity were evaluated using the paired Student's *t*-test, and P-values of < 0.05 were considered significant.

2.3 Results

2.3.1 Cloning of CFP-10/ESAT-6 Family Members

PCR amplification of the full length coding sequences of CFP-10 and ESAT-6 was analysed by gel electrophoresis. The results indicated that the PCR had been successful as DNA bands of the expected size (approximately 300 bp) were visualised in the agarose gel (Figure 2.4). Purified PCR products were ligated into the *GAL4* DBD vector (pGBD) and the *GAL4* TAD vector (pGAD). *E. coli* DH5 α cells transformed with the yeast two-hybrid constructs were selected by their ability to grow on LB media supplemented with ampicillin. Transformants were screened for the presence of CFP-10 and ESAT-6 inserts successfully ligated into the DBD and TAD vectors. Dual restricted (*EcoRI/BamHI*) plasmid samples were analysed by DNA gel electrophoresis. Results clearly show the presence of CFP-10 and ESAT-6 inserts in both pGBD and pGAD vectors (Figure 2.5). Integrity of the inserts was confirmed by DNA sequencing (PNACL) (Appendix 4.1- 4.4).

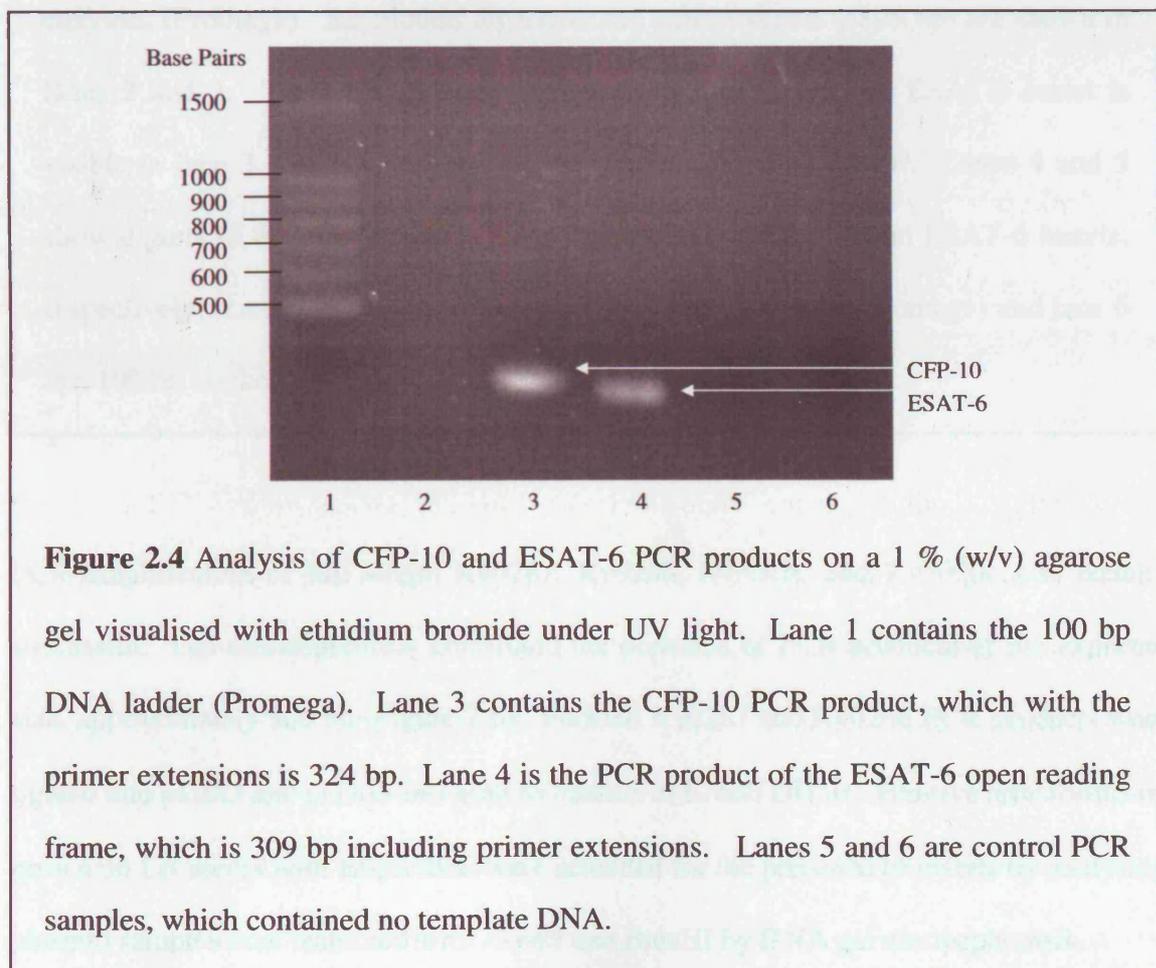


Figure 2.4 Analysis of CFP-10 and ESAT-6 PCR products on a 1 % (w/v) agarose gel visualised with ethidium bromide under UV light. Lane 1 contains the 100 bp DNA ladder (Promega). Lane 3 contains the CFP-10 PCR product, which with the primer extensions is 324 bp. Lane 4 is the PCR product of the ESAT-6 open reading frame, which is 309 bp including primer extensions. Lanes 5 and 6 are control PCR samples, which contained no template DNA.

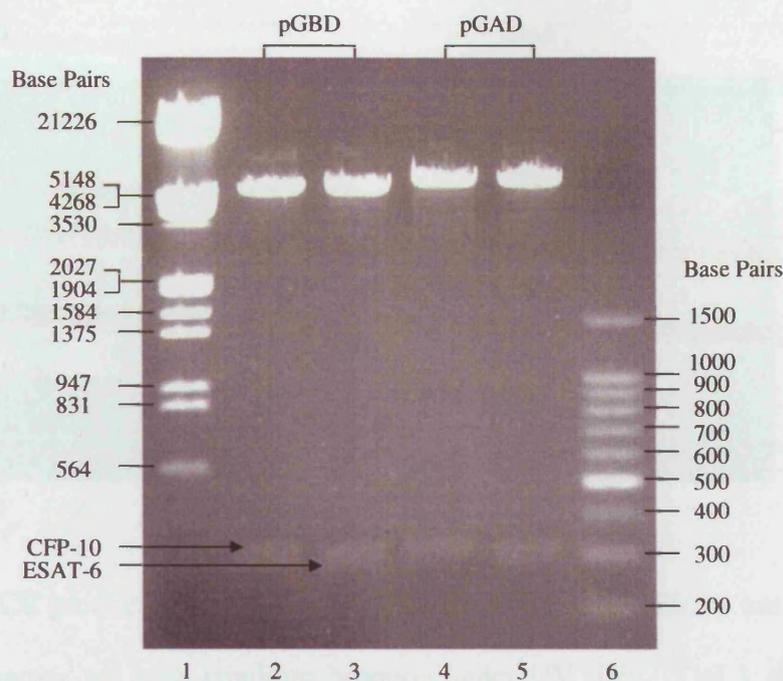


Figure 2.5 A 1 % (w/v) agarose gel showing dual restriction digests of successfully ligated yeast two hybrid vectors pGBD and pGAD containing the PCR inserts of CFP-10 and ESAT-6. The plasmids were digested with *EcoRI* and *BamHI* restriction enzymes (Promega). Restriction digests of the pGBD vector (5895 bp) are shown in lanes 2 and 3. The CFP-10 insert is present in lane 2, and the ESAT-6 insert is visible in lane 3. Both inserts run at the expected level in the gel. Lanes 4 and 5 show digests of the pGAD vector (6665 bp) containing CFP-10 and ESAT-6 inserts, respectively. Lane 1 contains λ DNA/*EcoRI* + *HindIII* marker (Promega) and lane 6 is a 100 bp marker (Promega).

PCR amplification of full length Rv0287, Rv0288, Rv3019c and Rv3020c was equally successful. Gel electrophoresis confirmed the presence of PCR products of the expected size, approximately 300 bp (Figure 2.6). Purified Rv0287 and Rv0288 PCR products were ligated into pGBD and pGAD and used to transform *E. coli* DH5 α . Positive transformants, grown on LB media with ampicillin, were screened for the presence of inserts by analysing plasmid samples dual restricted with *EcoRI* and *BamHI* by DNA gel electrophoresis.

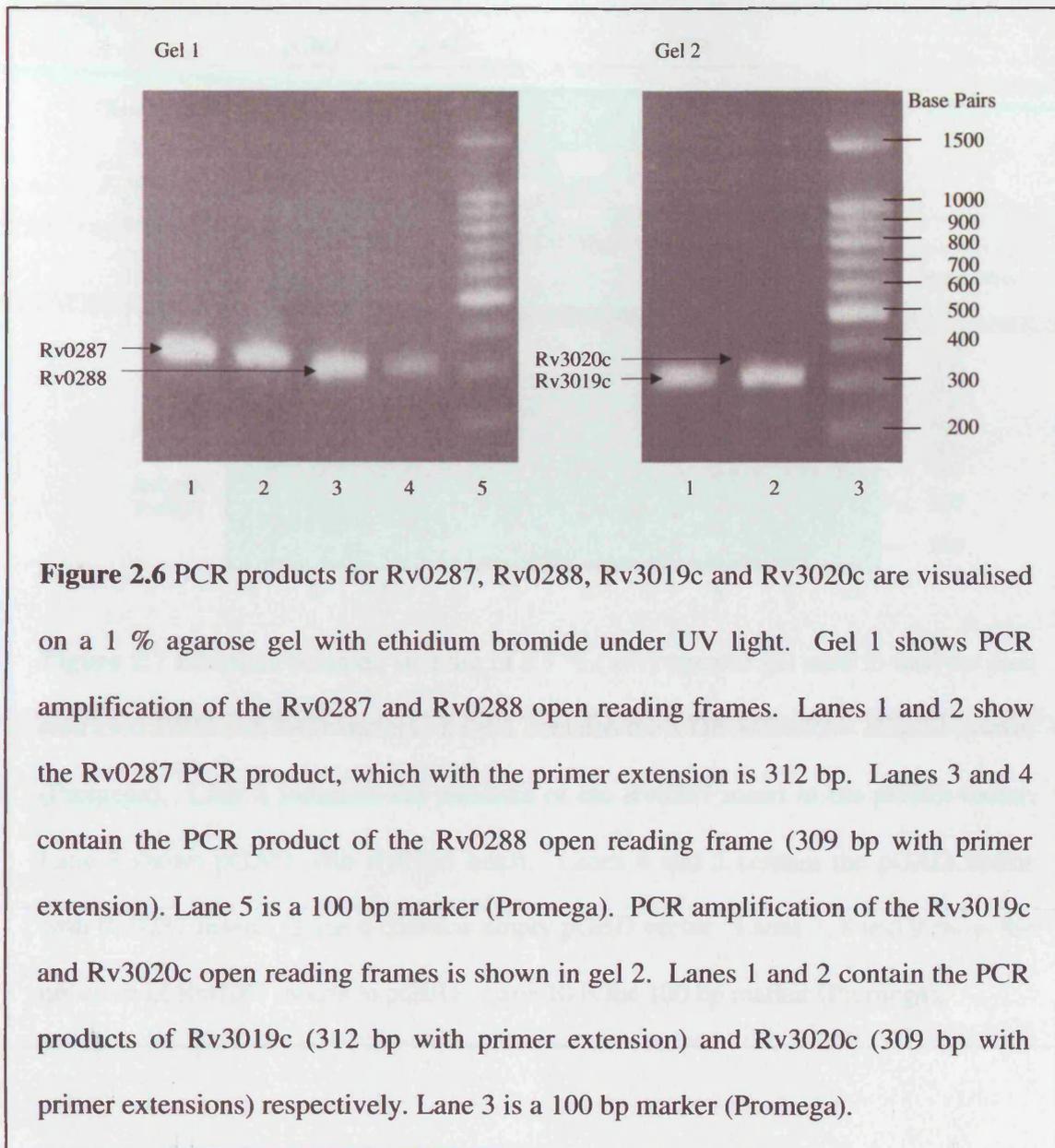


Figure 2.6 PCR products for Rv0287, Rv0288, Rv3019c and Rv3020c are visualised on a 1 % agarose gel with ethidium bromide under UV light. Gel 1 shows PCR amplification of the Rv0287 and Rv0288 open reading frames. Lanes 1 and 2 show the Rv0287 PCR product, which with the primer extension is 312 bp. Lanes 3 and 4 contain the PCR product of the Rv0288 open reading frame (309 bp with primer extension). Lane 5 is a 100 bp marker (Promega). PCR amplification of the Rv3019c and Rv3020c open reading frames is shown in gel 2. Lanes 1 and 2 contain the PCR products of Rv3019c (312 bp with primer extension) and Rv3020c (309 bp with primer extensions) respectively. Lane 3 is a 100 bp marker (Promega).

The gel shown in figure 2.7 demonstrates the successful ligation of the Rv0287 and Rv0288 PCR products into both the *GAL4* DBD and *GAL4* TAD vectors. Similar success was achieved for the ligation of Rv3019c and Rv3020c PCR products into pGBD and pGAD (results not shown). DNA sequencing was used to confirm the sequence and orientation of the PCR products in the pGBD and pGAD vectors (Appendix 4.5 - 4.12).

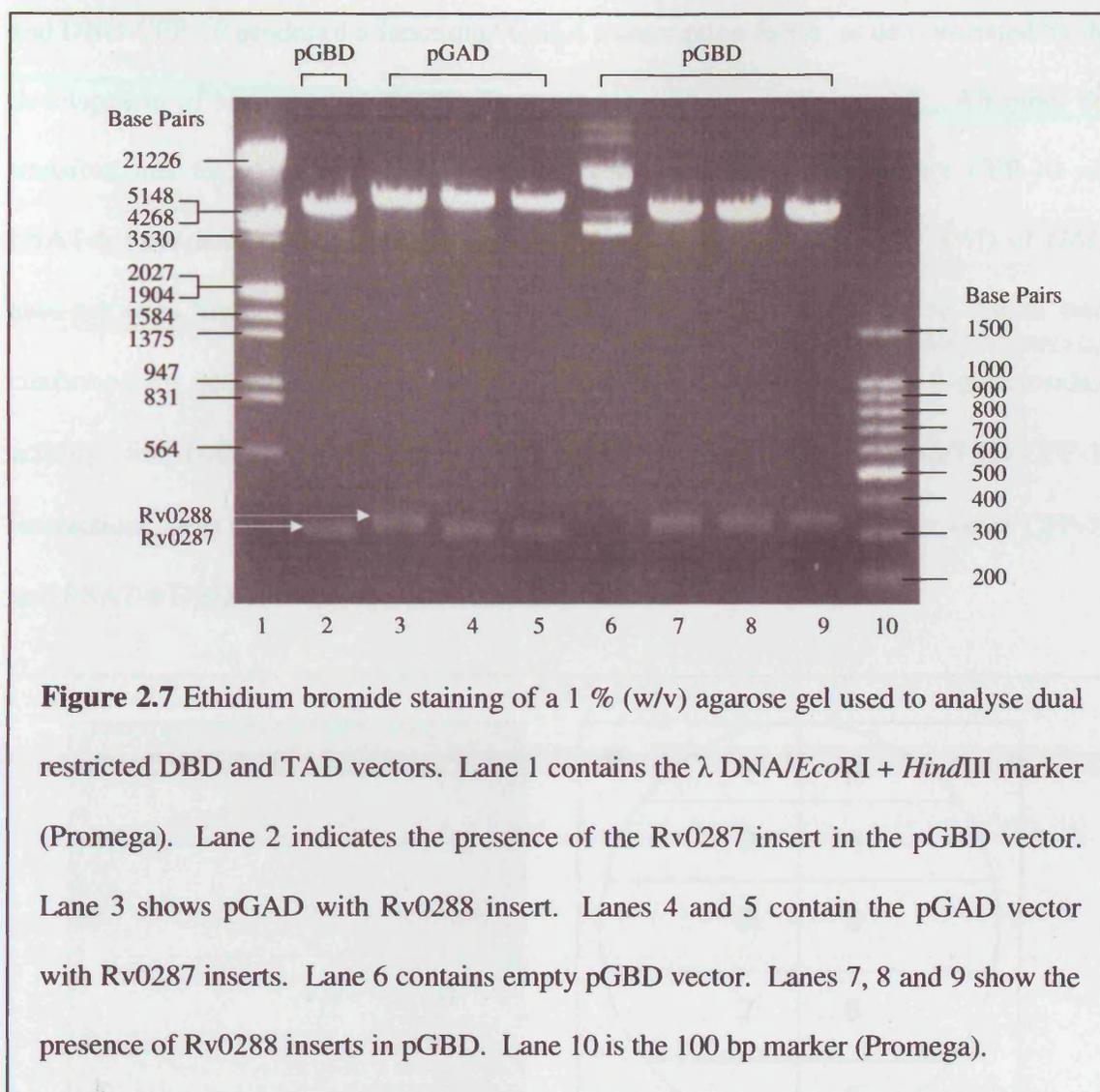


Figure 2.7 Ethidium bromide staining of a 1 % (w/v) agarose gel used to analyse dual restricted DBD and TAD vectors. Lane 1 contains the λ DNA/EcoRI + HindIII marker (Promega). Lane 2 indicates the presence of the Rv0287 insert in the pGBD vector. Lane 3 shows pGAD with Rv0288 insert. Lanes 4 and 5 contain the pGAD vector with Rv0287 inserts. Lane 6 contains empty pGBD vector. Lanes 7, 8 and 9 show the presence of Rv0288 inserts in pGBD. Lane 10 is the 100 bp marker (Promega).

2.3.2 Complex Formation between Genome Partners of the CFP-10/ESAT-6 Family

The genes encoding CFP-10, ESAT-6, Rv0287, Rv0288, Rv3019c and Rv3020c were successfully cloned into both pGBD and pGAD vectors, and were used in various combinations to transform *S. cerevisiae*. Initial yeast two hybrid studies were designed to determine whether this approach could detect the tight interactions seen *in vitro* for CFP-10 and ESAT-6. Yeast cells co-transformed with combinations of the *GAL4* DBD and *GAL4* TAD vectors, with and without CFP-10 and ESAT-6 fusions, were assayed for *lacZ* and *HIS3* reporter gene activity by filter lift assay and growth on histidine deficient media. Only yeast cells co-transformed with TAD-CFP-10 and DBD-ESAT-6 or TAD-ESAT-6

and DBD-CFP-10 produced a functional GAL4 transcription factor, as demonstrated by the development of blue colonies in the filter lift assay shown in figure 2.8. All other co-transformants tested in this assay remained white indicating that neither CFP-10 nor ESAT-6 form homodimers and that neither protein fused to the DBD or TAD of *GAL4* auto-activates transcription of the *lacZ* reporter gene (Figure 2.9). These results were confirmed by quantitative liquid assays (Figure 2.10 A) where specific β -galactosidase activity for both the DBD-CFP-10/TAD-ESAT-6 and DBD-ESAT-6/TAD-CFP-10 interactions were significantly above the β -galactosidase levels reported for other CFP-10 and ESAT-6 DBD/TAD combinations tested ($P < 0.05$).

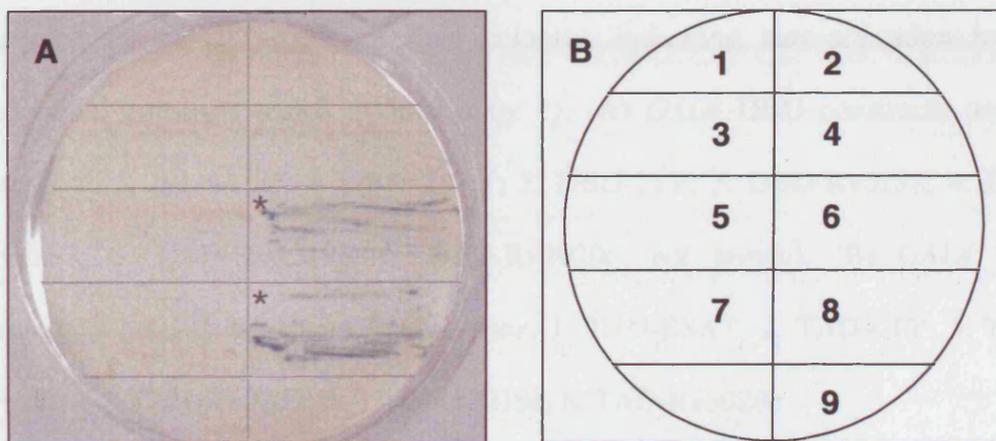


Figure 2.8 Detection of pair wise interactions between CFP-10 and ESAT-6. (A) Colony filter lift assay of yeast co-transformed with pGAD and pGBD constructs containing combinations of CFP-10 and ESAT-6 fused to either *GAL4* TAD or *GAL4* DBD. Expression of the *lacZ* reporter gene is detected by the development of blue colonies, indicative of complex formation (highlighted by *). (B) Identification of the *GAL4* DBD and *GAL4* TAD fusions used to transform *S. cerevisiae*. 1. DBD + TAD; 2. DBD + TAD-ESAT-6; 3. DBD + TAD-CFP-10; 4. DBD-ESAT-6 + TAD; 5. DBD-ESAT-6 + TAD-ESAT-6; 6. DBD-ESAT-6 + TAD-CFP-10; 7. DBD-CFP-10 + TAD; 8. DBD-CFP-10 + TAD-ESAT-6; 9. DBD- CFP-10 + TAD-CFP-10.

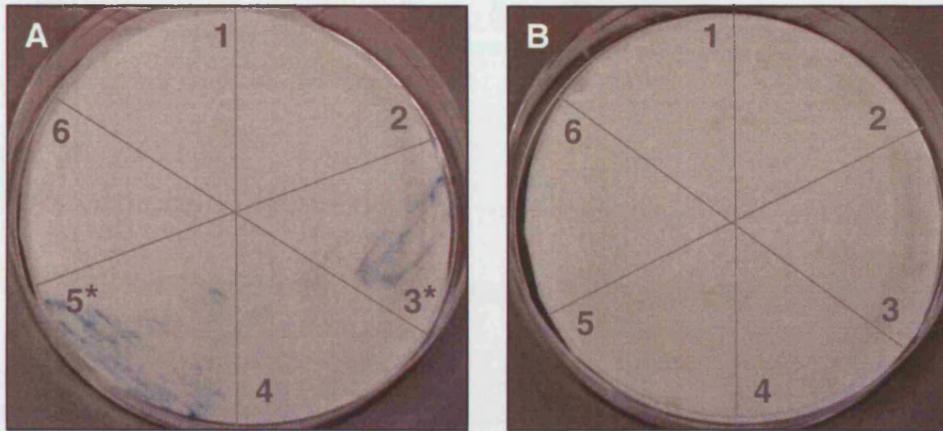


Figure 2.9 Filter lift assay of yeast transformed with (A) pGBD and (B) pGAD constructs containing ESAT-6, CFP-10, Rv0288, Rv0287, Rv3019c and Rv3020c fused to either *GAL4* DBD or *GAL4* TAD. Expression of the *lacZ* reporter gene is detected by the formation of blue colonies, indicating auto-activation by the individual construct tested (indicated by *). (A) *GAL4* DBD constructs used to transform *S. cerevisiae*. 1. DBD-ESAT; 2. DBD-CFP; 3. DBD-Rv0288; 4. DBD-Rv0287; 5. DBD-Rv3019c; 6. DBD-Rv3020c (not grown). (B) *GAL4* TAD constructs used to transform *S. cerevisiae*. 1. TAD-ESAT; 2. TAD-CFP; 3. TAD-Rv0288; 4. TAD-Rv0287; 5. TAD-Rv3019c; 6. TAD-Rv3020c.

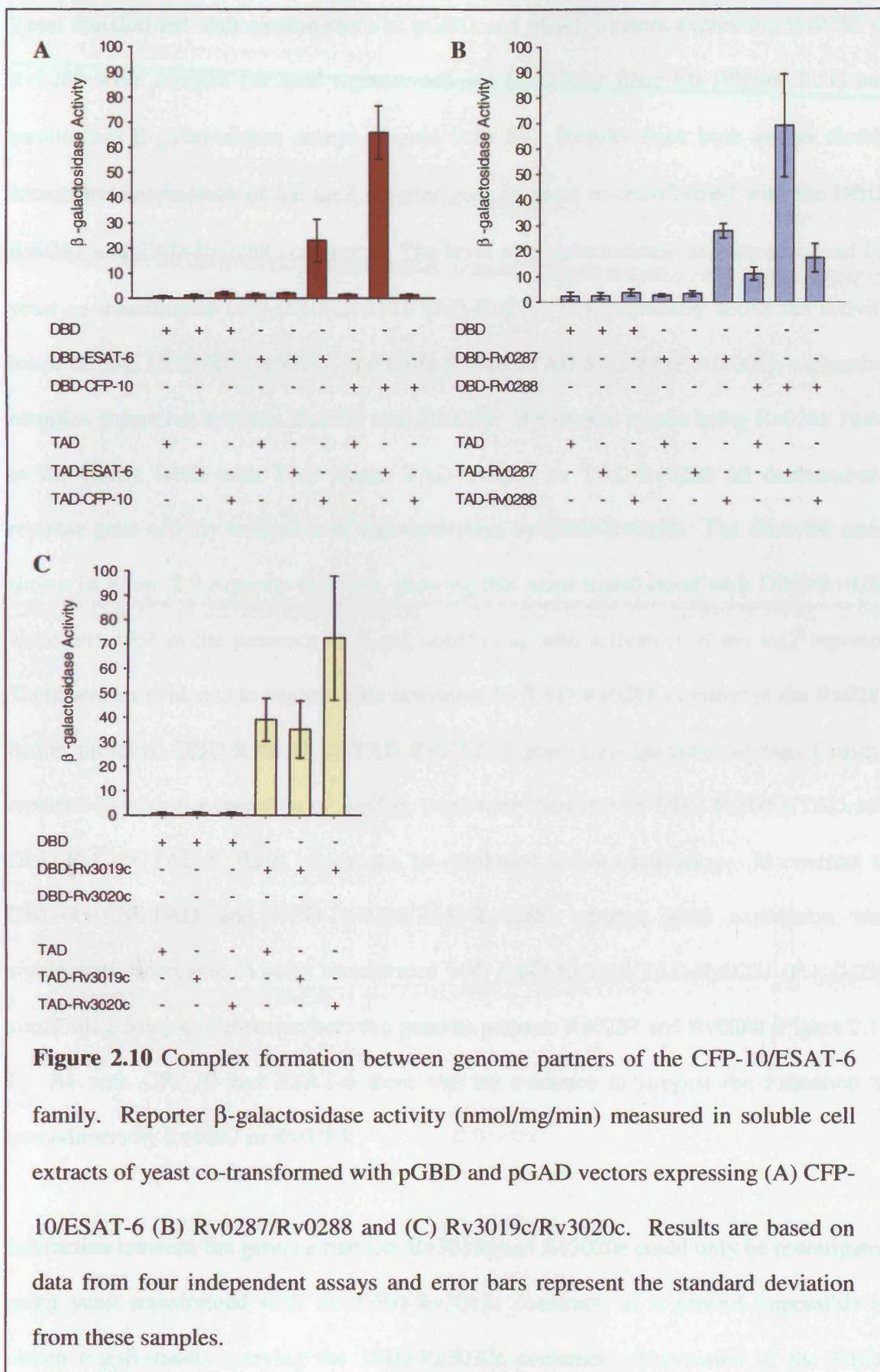


Figure 2.10 Complex formation between genome partners of the CFP-10/ESAT-6 family. Reporter β -galactosidase activity (nmol/mg/min) measured in soluble cell extracts of yeast co-transformed with pGBD and pGAD vectors expressing (A) CFP-10/ESAT-6 (B) Rv0287/Rv0288 and (C) Rv3019c/Rv3020c. Results are based on data from four independent assays and error bars represent the standard deviation from these samples.

Yeast transformed with combinations of pGBD and pGAD vectors expressing Rv0287 or Rv0288 were assayed for *lacZ* reporter activity by colony filter lift (Figure 2.11) and quantitative β -galactosidase assays (Figure 2.10 B). Results from both assays clearly demonstrate expression of the *lacZ* reporter gene in yeast co-transformed with the DBD-Rv0287 and TAD-Rv0288 constructs. The level of β -galactosidase activity observed for yeast co-transformed with DBD-Rv0287/TAD-Rv0288 is significantly above the activity levels of both DBD-Rv0287/TAD and DBD-Rv0287/TAD-Rv0287 ($P < 0.001$), suggesting complex formation between Rv0287 and Rv0288. Reciprocal assays using Rv0288 fused to the GAL4 DBD with TAD alone, TAD-Rv0287 or TAD-Rv0288 all demonstrated reporter gene activity indicative of auto-activation by DBD-Rv0288. The filter lift assay shown in figure 2.9 supports this data, showing that yeast transformed with DBD-Rv0288 alone turn blue in the presence of X-gal, confirming auto-activation of the *lacZ* reporter. There was no evidence to suggest auto-activation by TAD-Rv0288 or either of the Rv0287 fusion proteins, DBD-Rv0287 or TAD-Rv0287 (Figure 2.9). Quantitative liquid assays reported low level expression of *lacZ* in yeast transformed with DBD-Rv0288/TAD and DBD-Rv0288/TAD-Rv0288 which can be attributed to auto-activation. In contrast to DBD-Rv0288/TAD and DBD-Rv0288/TAD-Rv0288, reporter gene expression was significantly increased in yeast transformed with DBD-Rv0288/TAD-Rv0287 ($P < 0.05$); confirming complex formation between genome partners Rv0287 and Rv0288 (Figure 2.10 B). As with CFP-10 and ESAT-6 there was no evidence to suggest the formation of homodimers by Rv0287 or Rv0288.

Interaction between the genome partners Rv3019c and Rv3020c could only be investigated using yeast transformed with the DBD-Rv3019c construct, as it proved impossible to obtain transformants carrying the DBD-Rv3020c construct. Expression of the DBD-Rv3019c fusion protein alone was shown to induce auto-activation of the *lacZ* reporter

gene (Figure 2.9). However, β -galactosidase activity of yeast doubly transformed with DBD-Rv3019c and TAD-Rv3020c was considerably higher than that observed for DBD-Rv3019c with TAD alone or TAD-Rv3019c ($P < 0.05$), indicating that Rv3019c and Rv3020c dimerise to form stable complexes like the genome pairs CFP-10/ESAT-6 and Rv0287/Rv0288 (Figure 2.10 C).

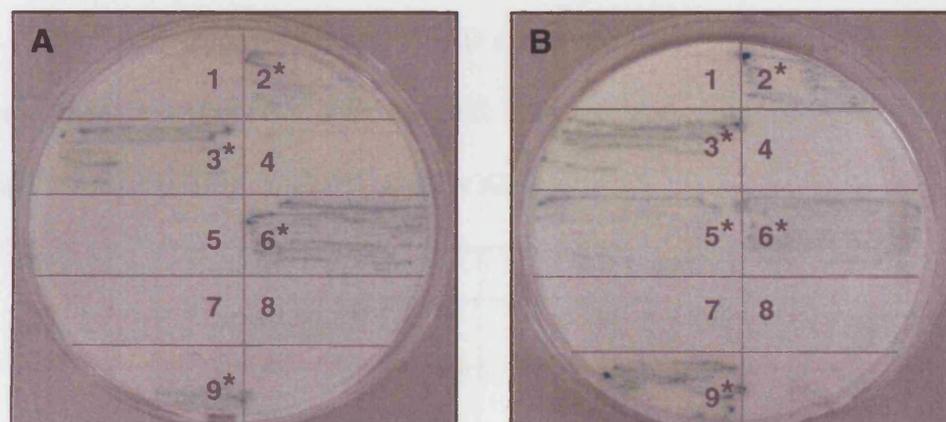


Figure 2.11 Yeast two-hybrid detection of interactions between members of the CFP-10/ESAT-6 family. Results from X-gal filter lift assays, where the development of blue colonies indicates *lacZ* reporter gene activity (positive interactions are highlighted by *). (A) Rv0287 fused to GAL4 DBD of GAL4 TAD is co-transformed with combinations of CFP-10, ESAT-6 and Rv0288. 1. DBD + TAD; 2. DBD-CFP-10 + TAD-ESAT-6; 3. DBD-ESAT-6 + TAD-CFP-10; 4. DBD-Rv0287 + TAD-ESAT-6; 5. DBD-Rv0287 + TAD-CFP-10; 6. DBD-Rv0287 + TAD-Rv0288; 7. DBD-ESAT-6 + TAD-Rv0287; 8. DBD-CFP-10 + TAD-Rv0287; 9. DBD-Rv0288 + TAD-Rv0287. (B) Yeast co-transformed with Rv0288 constructs and combinations of CFP-10, ESAT-6 and Rv0287. 1. DBD + TAD; 2. DBD-CFP-10 + TAD-ESAT-6; 3. DBD-ESAT-6 + TAD-CFP-10; 4. DBD-Rv0288 + TAD-ESAT-6; 5. DBD-Rv0288 + TAD-CFP-10; 6. DBD-Rv0288 + TAD-Rv0287; 7. DBD-ESAT-6 + TAD-Rv0288; 8. DBD-CFP-10 + TAD-Rv0288; 9. DBD-Rv0287 + TAD-Rv0288.

2.3.3 Interactions between Non-Genome Partners of the CFP-10/ESAT-6 Family

The yeast two-hybrid approach was also employed to investigate complex formation between non-genome partners within the CFP-10/ESAT-6 family. CFP-10 and ESAT-6 fused to *GAL4* DBD or *GAL4* TAD were screened against all possible combinations of Rv0287, Rv0288, Rv3019c and Rv3020c as described in table 2.1.

Table 2.1 Combinations of pGBD and pGAD constructs used to transform *S. cerevisiae* to investigate cross-talk between CFP-10 and ESAT-6 and with non-genome partners, including Rv0287, Rv0288, Rv3019c and Rv3020c.

pGBD Construct	pGAD Construct
pGBD-CFP-10	pGAD-Rv0287
pGBD-CFP-10	pGAD-Rv0288
pGBD-CFP-10	pGAD-Rv3019c
pGBD-CFP-10	pGAD-Rv3020c
pGBD-ESAT-6	pGAD-Rv0287
pGBD-ESAT-6	pGAD-Rv0288
pGBD-ESAT-6	pGAD-Rv3019c
pGBD-ESAT-6	pGAD-Rv3020c
pGBD-Rv0287	pGAD-CFP-10
pGBD-Rv0288	pGAD-CFP-10
pGBD-Rv3019c	pGAD-CFP-10
pGBD-Rv0287	pGAD-ESAT-6
pGBD-Rv0288	pGAD-ESAT-6
pGBD-Rv3019c	pGAD-ESAT-6

Colony filter lift assays, shown in figure 2.11, revealed that neither CFP-10 nor ESAT-6 interact with Rv0287 or Rv0288. The development of blue colonies in yeast doubly transformed with DBD-Rv0288 and TAD-CFP-10 is due to auto-activation by the DBD-Rv0288 fusion protein, as described in section 2.3.2. In the example shown, yeast did not grow when transformed with DBD-Rv0288 and TAD-ESAT-6, however these colonies

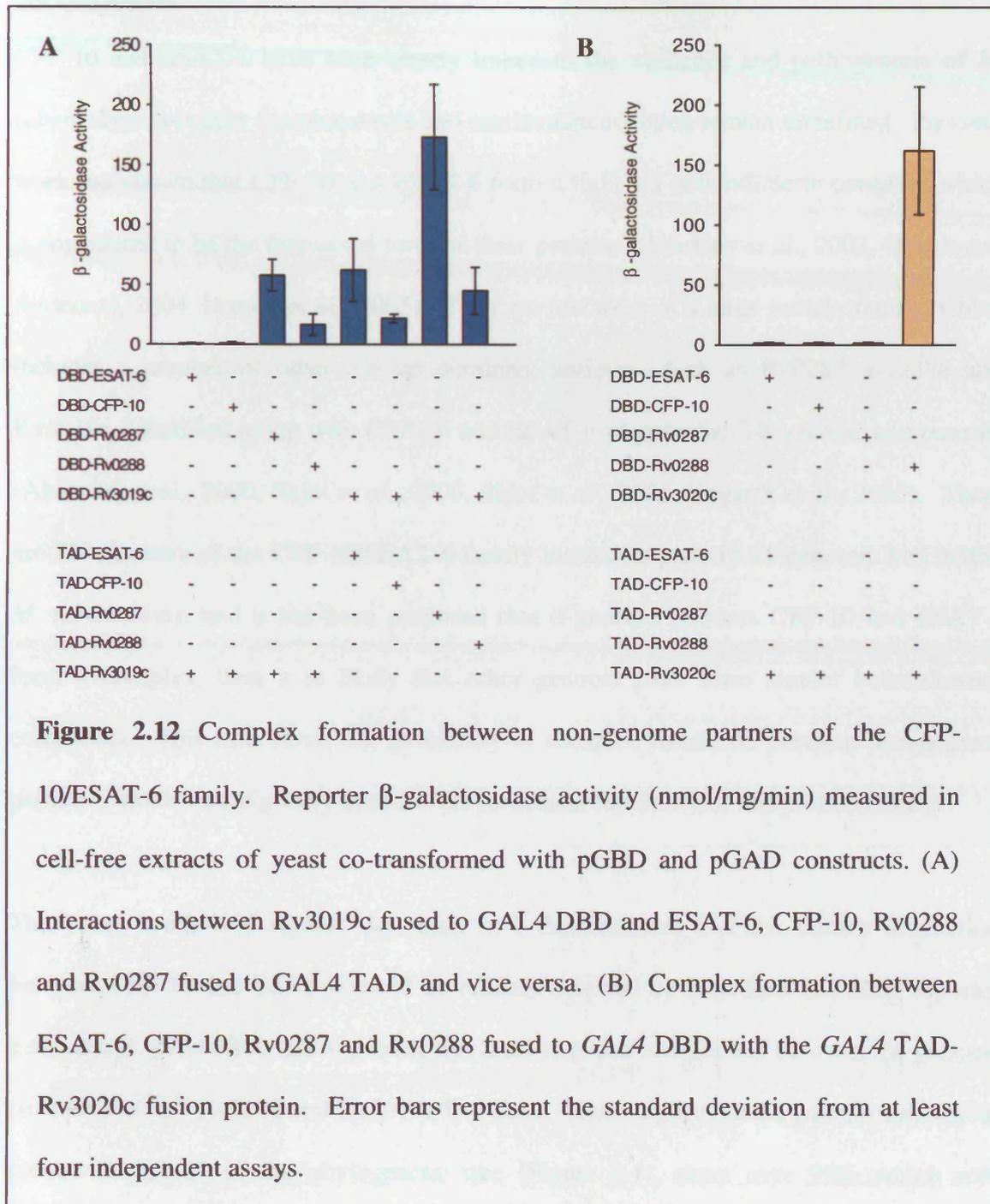
would also turn blue in the presence of X-gal as a result of auto-activation by DBD-Rv0288. Similarly, filter lifts (not shown) detected no interaction between CFP-10 and ESAT-6 with either Rv3019c or Rv3020c. These results show that neither CFP-10 nor ESAT-6 are involved in cross-talk with non-genome partners from Rv0287/Rv0288 or Rv3019c/Rv3020c, therefore it is suggested that CFP-10 and ESAT-6 are unlikely to interact with other members of the CFP-10/ESAT-6 protein family.

Cross-talk between Rv3019c and Rv3020c with other members of the CFP-10/ESAT-6 family were analysed by quantitative β -galactosidase assays (Figure 2.12). Rv3019c-based assays show no interaction between Rv3019c and CFP-10 or ESAT-6, confirming previous results from filter lifts (Figure 2.12 A). Quantitative liquid assays detected no reporter gene activity in yeast transformed with DBD-CFP-10 or DBD-ESAT-6 with TAD-Rv3019c. However, β -galactosidase activity reported in reciprocal assays was attributed to auto-activation of *lacZ* by the Rv3019c bait protein, as described previously (Figure 2.9). β -galactosidase activity observed in yeast co-transformed with DBD-Rv0288/TAD-Rv3019c and DBD-Rv3019c/TAD-Rv0288 was also due to auto-activation, as both Rv0288 and Rv3019c fused to the DBD can auto-activate transcription of the *lacZ* reporter gene.

Results from the β -galactosidase assays demonstrate that *lacZ* expression by DBD-Rv0287/TAD-Rv3019c is statistically significant in comparison to the base level of activity by DBD-ESAT-6/TAD-Rv3019c and DBD-CFP-10/TAD-Rv3019c ($P < 0.01$), and to the background level of auto-activation by DBD-Rv0288/TAD-Rv3019c ($P < 0.05$). In the reciprocal assays, where DBD-Rv3019c is responsible for auto-activation, the level of β -galactosidase activity observed for DBD-Rv3019c/TAD-Rv0287 is considerably above the level of auto-activation reported for DBD-Rv3019c/TAD-ESAT-6, DBD-

Rv3019c/TAD-CFP-10 and DBD-Rv3019c/TAD-Rv0288. Based on the results from *t*-tests the P values were calculated to be < 0.01, demonstrating statistical significance. In conclusion, β -galactosidase activity of both DBD-Rv3019c/TAD-Rv0287 and DBD-Rv0287/TAD-Rv3019c are significantly above the background level of activity for all other combinations tested when normalised for auto-activation activity ($P < 0.05$). Therefore, these results strongly suggest the formation of a heterodimeric complex between Rv0287 and Rv3019c.

Interactions with Rv3020c were only investigated using Rv3020c fused to GAL4 TAD, as it was not possible to transform yeast with the DBD-Rv3020c construct in any combination. As shown in figure 2.12 B, the β -galactosidase activity of yeast co-transformed with DBD-Rv0288 and TAD-Rv3020c is obviously above the activity levels observed for DBD-ESAT-6/TAD-Rv3020c, DBD-CFP-10/TAD-Rv3020c and DBD-Rv0287/TAD-Rv3020c ($P < 0.05$), clearly demonstrating that TAD-Rv3020c does not interact with CFP-10, ESAT-6 or Rv0287 fused to *GAL4* DBD. The reported β -galactosidase activity for DBD-Rv0288/TAD-Rv3020c is clearly above the background level of auto-activation observed for other DBD-Rv0288 combinations (DBD-Rv0288/TAD, DBD-Rv0288/TAD-Rv0288 and DBD-Rv0288/TAD-3019c, $P < 0.05$), signifying cross-talk between non-genome partners Rv3020c and Rv0288.



2.4 Discussion

CFP-10 and ESAT-6 have been clearly linked to the virulence and pathogenesis of *M. tuberculosis*, yet their functional role and mechanism of action remain undefined. Previous work has shown that CFP-10 and ESAT-6 form a tight 1:1 heterodimeric complex, which is considered to be the functional form of these proteins (Renshaw *et al.*, 2002, Okkels and Anderson, 2004, Brodin *et al.*, 2005). They are members of a large protein family, which includes a number of other immunodominant antigens, such as Rv0287, Rv0288 and Rv3019c, identified along with CFP-10 and ESAT-6 as potential TB vaccine components (Alderson *et al.*, 2000, Skjot *et al.*, 2000, Skjot *et al.*, 2002, Hogarth *et al.*, 2005). There are 23 members of the CFP-10/ESAT-6 family located in pairs in 11 genomic loci within *M. tuberculosis*, and it has been proposed that if genome partners CFP-10 and ESAT-6 form a complex, then it is likely that other genome pairs form similar heterodimeric complexes. This also raises the possibility of complex formation between non-genome partners, which would greatly increase the functional flexibility of this protein family.

The yeast two-hybrid studies described here demonstrated that the known interaction between CFP-10 and ESAT-6 could be reliably detected by both *lacZ* and *HIS3* reporter gene assays. Therefore this approach was used to probe interactions between the genome partners Rv0287/Rv0288 and Rv3019c/Rv3020c. These highly related pairs, found within groups B and B' of the phylogenetic tree (Figure 2.1), share over 95% amino acid homology, and were therefore considered likely candidates for cross talk between non-genome partners. β -galactosidase assays clearly identified pair wise interactions between Rv0287/Rv0288 and Rv3019c/Rv3020c. These results were confirmed by fluorescence studies carried out by collaborators at the University of Kent and subsequently at the University of Leicester, demonstrating that the gene products of Rv0287 and Rv0288 form a 1:1 heterodimeric complex with properties similar to that of the CFP-10•ESAT-6

complex described by Renshaw *et al.* (2002). Furthermore, Okkels and Anderson (2004) also observed protein-protein interactions between genome pairs CFP-10/ESAT-6 and Rv0287/Rv0288 using Western-Western blotting and protein-print overlay techniques. These studies suggest that all genome pairs from the CFP-10/ESAT-6 protein family will form 1:1 heterodimeric complexes as described for CFP-10/ESAT-6.

Helical wheel representations showing the heptad repeats (**a-b-c-d-e-f-g**) of the CFP-10 and ESAT-6 α -helices (Figure 2.13) and amino acid conservation between members of the CFP-10/ESAT-6 family (Figure 2.14) support the proposal that CFP-10/ESAT-6 complexes are likely to adopt a four-helix bundle structure similar to that reported for the CFP-10•ESAT-6 complex, as shown in figure 2.15 (Renshaw *et al.*, 2005).

The helical wheel representations shown in figure 2.13 are as expected for four helix-bundle; where the residues at positions **a** and **d** form the hydrophobic faces of the α -helices. Residues at these positions are buried within the core of the protein complex and are shielded from the solvent by polar and charged residues at positions **b**, **c**, **e** and **g** (Betz *et al.*, 1997). Analysis of the structure of the CFP-10•ESAT-6 complex reveals that residues at positions **a**, **e**, and occasionally **b** form the intramolecular interface between two helices in CFP-10 (Blue, top) and ESAT-6 (Blue, bottom), those at positions **c** and **g** form the intermolecular interface between CFP-10 and ESAT-6 (Red) and residues at position **d** are generally involved at both interfaces (Green).

Highly conserved residues were identified by multiple sequence alignments of the *M. tuberculosis* CFP-10/ESAT-6 protein family, in which Rv1037c and Rv1038c were included as representatives of the closely related A and A' families (Figure 2.14). Highly conserved amino acids are highlighted by asterisks (*) (Figure 2.13), clearly indicating that the residues at positions **a** and **d** are highly conserved throughout the family, suggesting

that these key interface residues stabilise the formation of the helix-turn-helix structures of the individual family members and suggest that the pairs are likely to adopt an four helix-bundle structure similar to that of CFP-10•ESAT-6 (Figure 2.15). Interface residues at positions **b/e** and **c/g** are poorly conserved, which suggests that they contribute to the stability and specificity of the protein complexes, therefore alteration of these amino acids, particularly at positions **g** and also **c** (at the intermolecular interface) will have an effect on complex stability and specificity. Residues at positions **g** and **c**, particularly within the N-terminal helices, are not conserved between CFP-10/ESAT-6, Rv0287/Rv0288 and Rv3019c/Rv3020c proteins indicating that cross-talk between these family members is unlikely. However, there is a high level of similarity between Rv0287/Rv0288 and Rv3019c/Rv3020c, which would predict favourable interactions between Rv0287/Rv3019c and Rv0288/Rv3020c. Similarly the sequence similarity within the A and A' families suggest that these family members will be able to cross-talk with non-genome pairs within the A/A' group (Figure 2.1).

Salt bridges between (Glu₁₄-Lys₃₈ and Glu₇₁-Lys₅₇) stabilise the interactions between CFP-10 and ESAT-6. These residues are not conserved throughout the family and are therefore unlikely to determine which pairs form complexes (Renshaw *et al.*, 2005). However, based on evidence from the helical wheels (Figure 2.13) and the structure of the CFP-10/ESAT-6 complex it is possible that salt bridges may be formed at a different position between Rv0287/Rv0288 and Rv3019c/Rv3020c (Lys₂₁-Glu/Asp₃₁ and Lys₆₄/Asp₆₄).

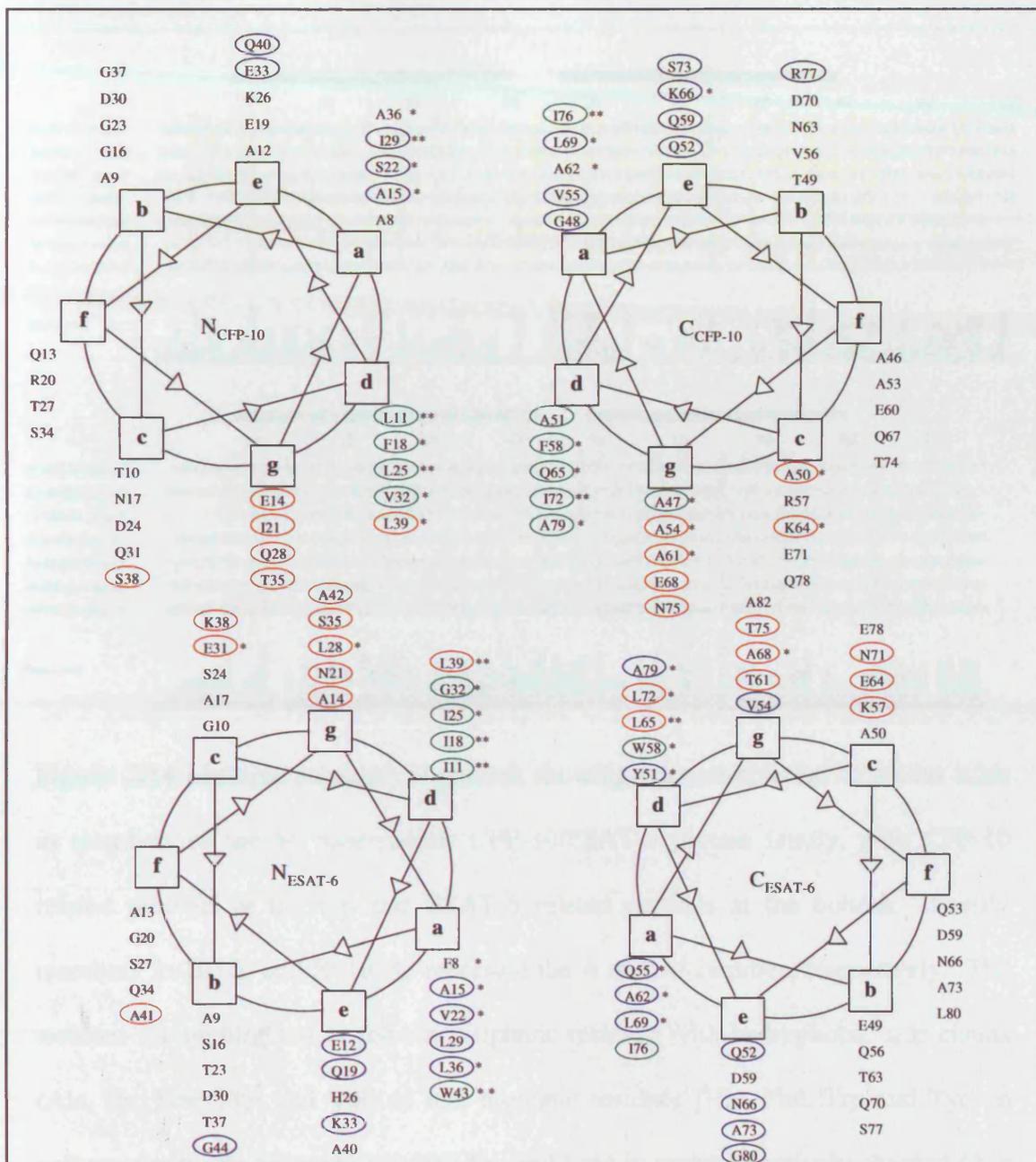


Figure 2.13 Helical wheel representations of CFP-10 (top) and ESAT-6 (bottom).

Residues at positions **a** and **d** form the hydrophobic core of the protein complex.

Residues highlighted in blue indicate interactions between two helices of CFP-10 or

ESAT-6, and those in red are the residues forming the intermolecular interface between

CFP-10 and ESAT-6. Residues highlighted in green are involved in both interfaces.

Based on the amino acids groups used in the multiple sequence alignments (Figure 2.14),

residues which are highly conserved (in the same group) throughout all of the CFP-

10/ESAT-6 family members are highlighted by two asterisks (**) and those conserved

through the majority of family members are highlighted by one asterisk (*)

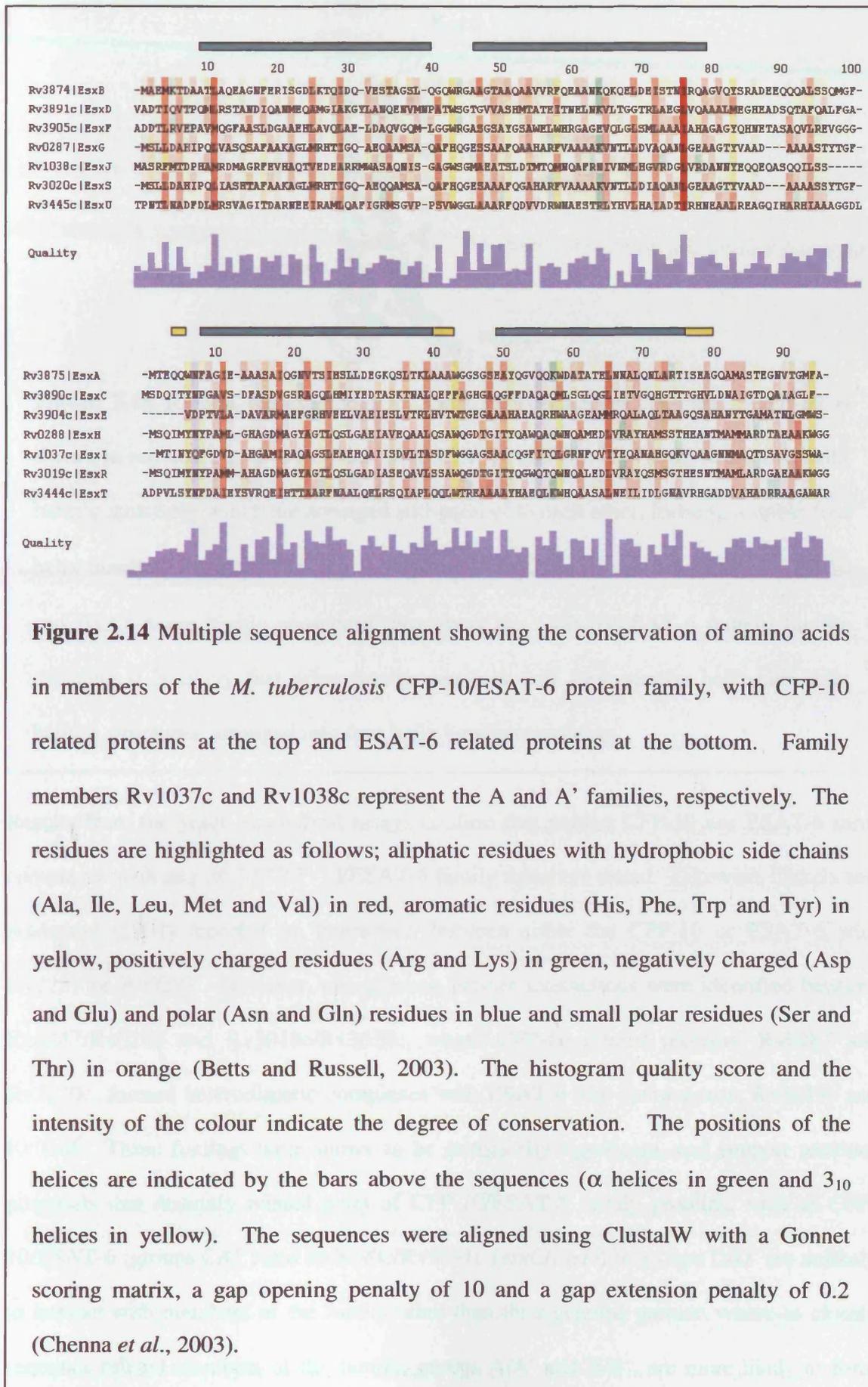


Figure 2.14 Multiple sequence alignment showing the conservation of amino acids in members of the *M. tuberculosis* CFP-10/ESAT-6 protein family, with CFP-10 related proteins at the top and ESAT-6 related proteins at the bottom. Family members Rv1037c and Rv1038c represent the A and A' families, respectively. The residues are highlighted as follows; aliphatic residues with hydrophobic side chains (Ala, Ile, Leu, Met and Val) in red, aromatic residues (His, Phe, Trp and Tyr) in yellow, positively charged residues (Arg and Lys) in green, negatively charged (Asp and Glu) and polar (Asn and Gln) residues in blue and small polar residues (Ser and Thr) in orange (Betts and Russell, 2003). The histogram quality score and the intensity of the colour indicate the degree of conservation. The positions of the helices are indicated by the bars above the sequences (α helices in green and 3_{10} helices in yellow). The sequences were aligned using ClustalW with a Gonnet scoring matrix, a gap opening penalty of 10 and a gap extension penalty of 0.2 (Chenna *et al.*, 2003).

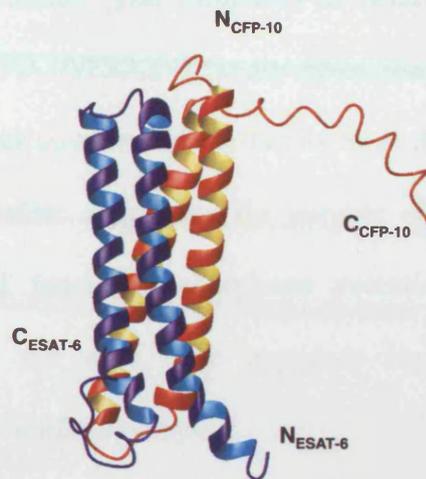


Figure 2.15 Ribbon representation of the CFP-10•ESAT-6 complex. CFP-10 is shown in red and ESAT-6 in blue. Both CFP-10 and ESAT-6 form helix-turn-helix hairpin structures which are arranged anti-parallel to each other, forming a stable four helix-bundle. Residues forming the hydrophobic faces of the α -helices of CFP-10 and ESAT-6 are highly conserved throughout the CFP-10/ESAT-6 protein family, therefore it is likely that other family members will form similar helix-turn-helix hairpin structures, arranged into four helix bundle complexes.

Results from the yeast two-hybrid assays confirm that neither CFP-10 nor ESAT-6 form complexes with any of the CFP-10/ESAT-6 family members tested. Likewise, Okkels and Anderson (2004) reported no interaction between either the CFP-10 or ESAT-6 with Rv0287 or Rv0288. However, non-genome partner interactions were identified between Rv0287/Rv0288 and Rv3019c/Rv3020c, where CFP-10 related proteins, Rv0287 and Rv3020c, formed heterodimeric complexes with ESAT-6 like counterparts, Rv3019c and Rv0288. These findings were shown to be statistically significant, and support previous proposals that distantly related pairs of CFP-10/ESAT-6 family proteins, such as CFP-10/ESAT-6 (groups C/C') and Rv3890c/Rv3891c (*esxC/esxD*) in groups D/D' are unlikely to interact with members of the family other than their genome partner, whereas closely sequence related members of the family, groups A/A' and B/B', are more likely to form

non-genome partner complexes. The formation of heterodimeric complexes between genome partners of the CFP-10/ESAT-6 family gives rise to 11 functional complexes. However, more promiscuous members of the family have the ability to form dimers with non-genome partners, therefore amplifying the number of CFP-10/ESAT-6 complexes. Increasing the number of functional complexes potentially increases the functional flexibility of the family and may have important implications for virulence and pathogenicity of the *M. tuberculosis* complex.

A recent study has shown that individual CFP-10 and ESAT-6 proteins induce a different T-cell response (IFN γ and TNF α) in comparison to the CFP-10•ESAT-6 complex (Marei *et al.*, 2005). The isolated individual CFP-10 and ESAT-6 proteins were readily digested by the lysosomal enzymes cathepsin L and S, whereas under the same conditions the CFP-10•ESAT-6 is resistant, suggesting that the complex is protected from antigen processing. Various members of the CFP-10/ESAT-6 family also induce individual T-cell responses (Alderson *et al.*, 2000, Skjot *et al.*, 2000, Skjot *et al.*, 2002), however it is unknown whether processing of other CFP-10/ESAT-6 family complexes, such as Rv0287•Rv0288 or Rv3019c•Rv3020c, is similar to that of CFP-10/ESAT-6. Interestingly, both chemical and heat denaturation studies have already shown that the Rv0287•Rv0288 complex is more stable than the CFP-10•ESAT-6 complex, with midpoints for guanidine hydrochloride denaturation and thermal denaturation of 0.9M/1.8M and 50/70 °C respectively (Personal communication – P. Renshaw, Department of Biochemistry, University of Leicester). Therefore it is quite likely that the Rv0287/Rv0288 complex is processed differently to CFP-10•ESAT-6.

The genes flanking CFP-10 and ESAT-6 encode a specialised secretion system for the active export of the CFP-10•ESAT-6 complex (Tekaiia *et al.*, 1999, Hsu *et al.*, 2003, Pym

et al., 2003, Stanley et al., 2003, Guinn et al., 2004). Rv0287 and Rv0288 are found within a gene cluster similar to that observed for CFP-10/ESAT-6, where neighbouring genes include homologues of the proteins involved in the formation of the CFP-10/ESAT-6 secretory complex (Figure 2.16). In contrast, the gene cluster for Rv3019c/Rv3020c lacks the genes encoding such secretion apparatus, yet Rv3019c has been detected in the culture supernatant of *M. tuberculosis*. It is possible that due to the close sequence similarity between the Rv0287•Rv0288 and Rv3019c•Rv3020c complexes, the Rv0287/Rv0288 secretion system may allow secretion of the Rv3019c•Rv3020c heterodimer.

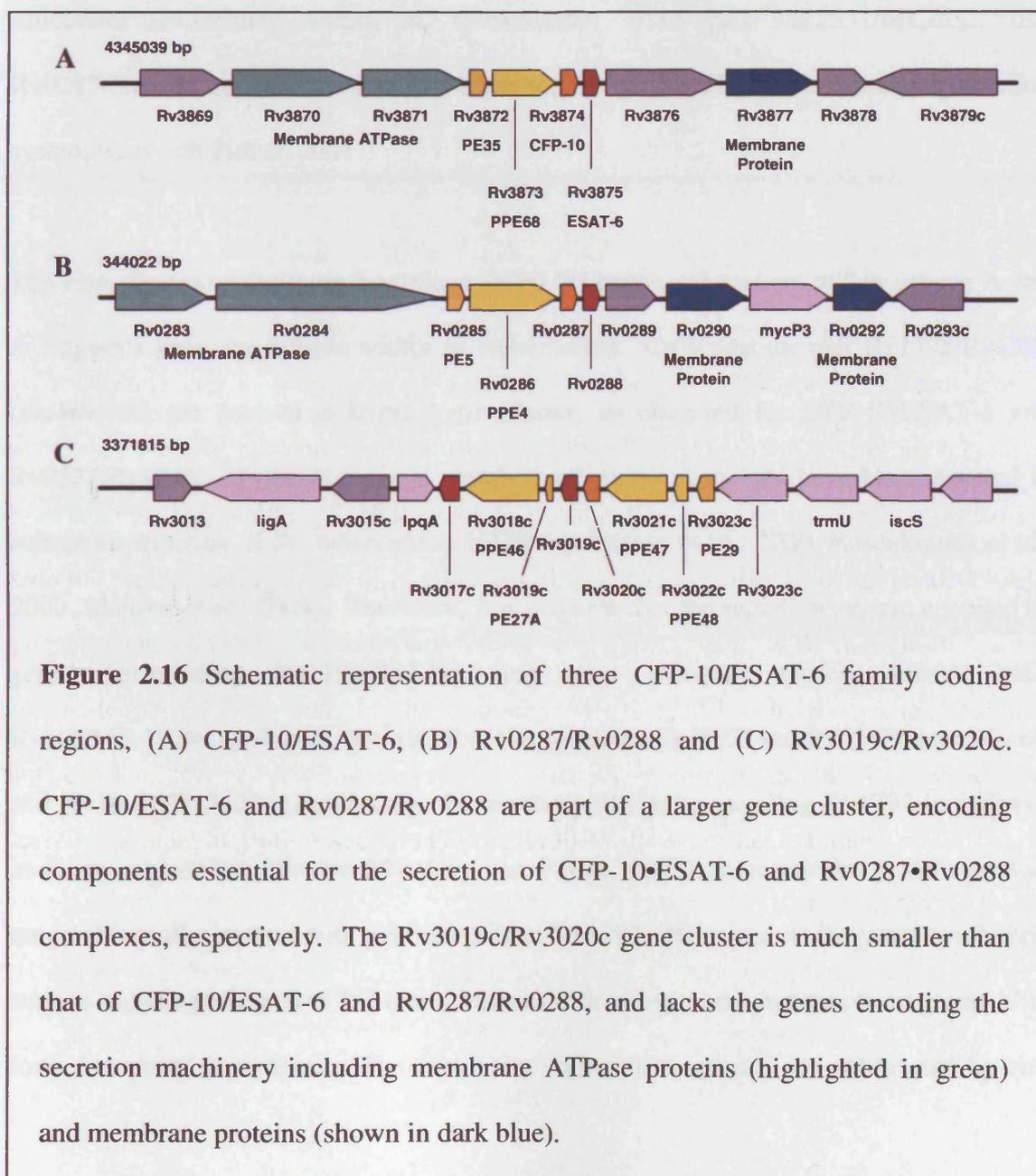


Figure 2.16 Schematic representation of three CFP-10/ESAT-6 family coding regions, (A) CFP-10/ESAT-6, (B) Rv0287/Rv0288 and (C) Rv3019c/Rv3020c. CFP-10/ESAT-6 and Rv0287/Rv0288 are part of a larger gene cluster, encoding components essential for the secretion of CFP-10•ESAT-6 and Rv0287•Rv0288 complexes, respectively. The Rv3019c/Rv3020c gene cluster is much smaller than that of CFP-10/ESAT-6 and Rv0287/Rv0288, and lacks the genes encoding the secretion machinery including membrane ATPase proteins (highlighted in green) and membrane proteins (shown in dark blue).

M. leprae has undergone reductive evolution, and is thought to contain the minimum gene set required to remain pathogenic (Cole *et al.*, 2001). Therefore, retention of CFP-10/ESAT-6 family proteins within *M. leprae* indicates the importance of this protein family in mycobacterial pathogens. CFP-10 and ESAT-6 are individually conserved within the *M. leprae* genome (ML0050 and ML0049 respectively), signifying their importance. As in *M. tuberculosis*, these genes are retained within a larger gene cluster encoding the specialised secretion system. Interestingly, Rv0287/Rv0288 and Rv3019c/Rv3020c are conserved as just a single pair within *M. leprae* suggesting functional redundancy within *M. tuberculosis*. The pair ML2531/ML2532, like Rv0287/Rv0288 are conserved within a larger gene cluster which also encodes a secretion system (Gey van Pittius, 2001).

The high degree of sequence homology (> 90 %) between members within groups A and A' suggests gene duplication within *M. tuberculosis*. Only genome pair Rv1792/Rv1793 (*esxM/esxN*) are part of a larger gene cluster, as observed for CFP-10/ESAT-6 and Rv0287/Rv0288. However, several members of groups A and A' have been detected in culture supernatant of *M. tuberculosis* H37Rv (Alderson *et al.*, 2000, Rosenkrands *et al.*, 2000, Mattow *et al.*, 2003). Therefore, it is possible that the secretion system encoded by genes surrounding Rv1792/Rv1793 may also transport closely related pairs Rv1037/Rv1038 (*esxI/esxJ*), Rv1197/Rv1198 (*esxK/esxL*), Rv2346c/Rv2347c (*esxO/esxP*) and Rv3619c/Rv3620c (*esxR/esxS*). Interestingly, the gene encoding Rv1792 includes an in-frame stop codon (residue 59) (Gey van Pittius, 2001), therefore the protein product may not actually interact with genome partner Rv1793. However, as the yeast two-hybrid studies suggest perhaps Rv1793 could interact with non-genome partners from group A' to form functional complexes. The five pairs within these groups are substituted by two

identical gene pairs within *M. leprae* (ML1055/ML1056 and ML1180/ML1181), which may suggest functional redundancy within *M. tuberculosis*.

In conclusion, results presented in this chapter strongly suggest that all genome pairs within the CFP-10/ESAT-6 family will form stable heterodimeric complexes. The yeast two hybrid studies also demonstrate that while closely related members of the CFP-10/ESAT-6 family (groups A/A' and B/B') will interact with non-genome partners, more distantly related members of the family, including CFP-10/ESAT-6 and Rv3890c/Rv3891c, are more likely to be restricted to complex formation with their genome partner. Finally, the helical wheel representations indicate that the residues found within the hydrophobic core of the protein complex are highly conserved throughout the CFP-10/ESAT-6 family, suggesting that the protein complexes are likely to form four helix-bundle structures similar to the CFP-10•ESAT-6 complex. Also specificity in complex formation is likely to be due to the residues located at the intermolecular interface. These residues are poorly conserved throughout the family therefore it is predicted that alterations of those amino acids will affect complex specificity and stability.

Chapter 3

Specific Binding of the CFP-10•ESAT-6 Complex to the Surface of Host Cells

3.1 Introduction

The *M. tuberculosis* secreted proteins CFP-10 and ESAT-6 play an essential role in virulence and pathogenesis, yet their function is still to be defined. The recently published solution structure of the CFP-10•ESAT-6 complex indicates that both proteins adopt very similar helix-turn-helix hairpin structures, arranged anti-parallel to each to form a stable four helix bundle, as shown in figure 3.1 (Renshaw *et al.*, 2005). It has previously been suggested that CFP-10 and ESAT-6 may form membrane pores, resulting in the lysis of host cells (Hsu *et al.*, 2003). However, the surface features of the complex and the studies presented in this chapter strongly argue against this type of role and are more consistent with a function based on binding to one, or more, target host cell proteins.

A recent microarray study revealed that the expression of CFP-10 and ESAT-6 are down-regulated by *M. tuberculosis* cells internalised within the macrophage phagosome, implying that the principal function of the complex is likely to occur prior to uptake of the bacterium by the macrophage (Schnappinger *et al.*, 2003). This study, together with the structural features of the complex, suggests that the CFP-10•ESAT-6 complex may bind to a receptor on the surface of host cells. The presence of a specialised secretion system for the ATP-dependant export of CFP-10 and ESAT-6 also suggests that these proteins play an important functional role following their release from the mycobacterial cell.

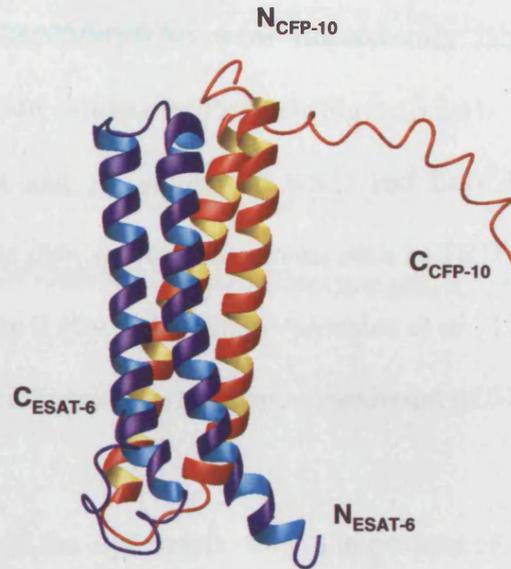
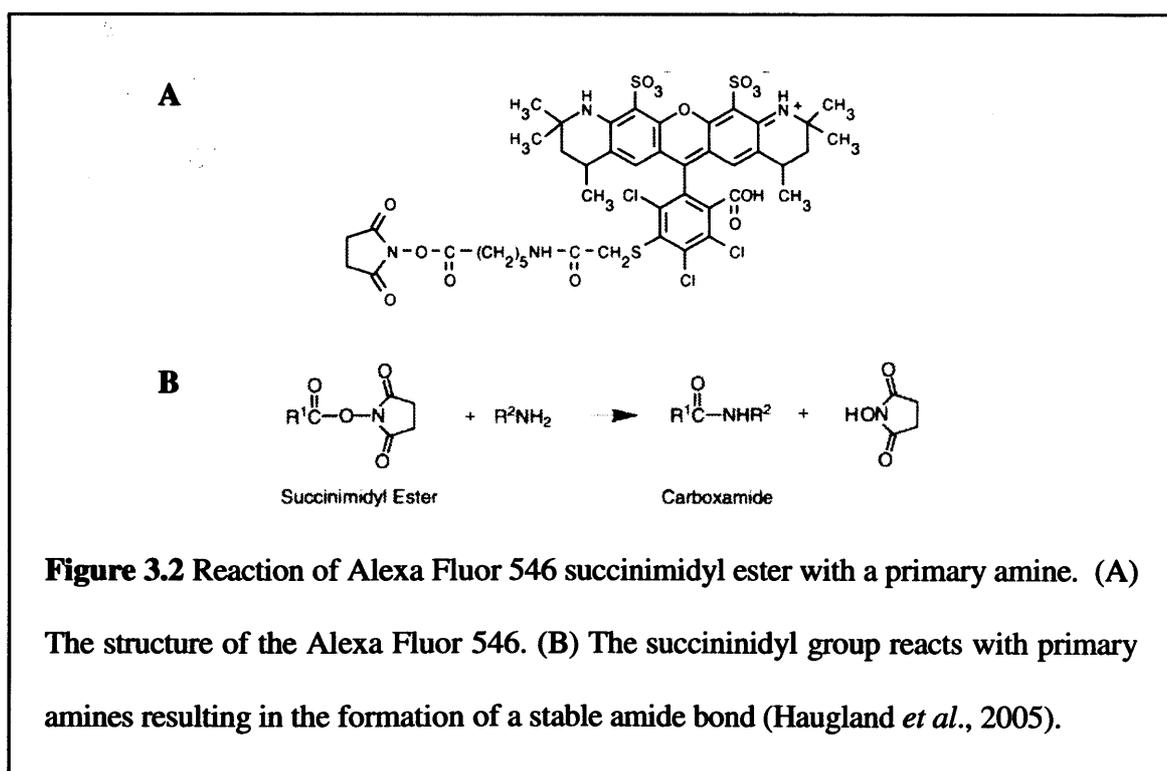


Figure 3.1 Ribbon representation of the solution structure of the CFP-10•ESAT-6 complex (Renshaw *et al.*, 2005). CFP-10 (red) and ESAT-6 (blue) both form helix-turn-helix hairpin structures, which lie anti-parallel to each other forming a stable four helix bundle. The C-terminal residues of CFP-10 clearly show the propensity to adopt a helical conformation, and along with the C-terminal of ESAT-6 they form flexible arms at each end of the four-helix bundle.

To test the hypothesis that host cells possess a specific cell surface receptor for the CFP-10•ESAT-6 complex, a fluorescence microscopy approach was used. Experiments were carried out with a range of cell types, including primary monocytes and monocyte-derived macrophages, human monocytic cell lines, MonoMac6 (MM6) (Ziegler-Heitbrock, 1988) and U937 (Sundstrom and Nilsson, 1976) and fibroblast cell lines, NIH-3T3 (Jainchill *et al.*, 1969) and Cos-1 (Gluzman, 1981). Human macrophages are the primary target for infection by *M. tuberculosis*, and we would therefore expect these types of cells to express a putative cell surface receptor for the CFP-10•ESAT-6 complex.

Samples of the CFP-10•ESAT-6 complex and an unrelated control protein from the *M. tuberculosis* complex (MPT70/MPB70) were fluorescently labelled using the amine-reactive dye Alexa Fluor 546 (Molecular Probes) (Figure 3.2A). Alexa Fluor dyes are pH insensitive between pH 4 and 10, soluble in water and have been shown to be more fluorescent and photostable than spectral analogues, such as TRITC (tetramethylrhodamine isothiocyanate) and cyanine 3 (Cy3) (Panchuk-Voloshina *et al.*, 1999). Alexa Fluor 546 is excited at 546 nm and has a fluorescence emission maximum of 573 nm.

The succinimidyl group of the dye reacts with non-protonated aliphatic amine groups, including the amino-terminal of proteins and the ϵ -amino groups of lysine residues (Figure 3.3). The N-terminal of the protein can be selectively labelled at around neutral pH as the N-terminal amine has a significantly lower pK_a (~ 8.0) than the lysine ϵ -amino groups (~ 10.0), which are optimally modified at pH 8.5 – 9.5 (Haugland *et al.*, 2005). Therefore, at pH 7.5 we would expect that the succinimidyl group of Alexa Fluor 546 would react with the primary amines of CFP-10 and ESAT-6, but not the ϵ -amino groups of lysine residues.

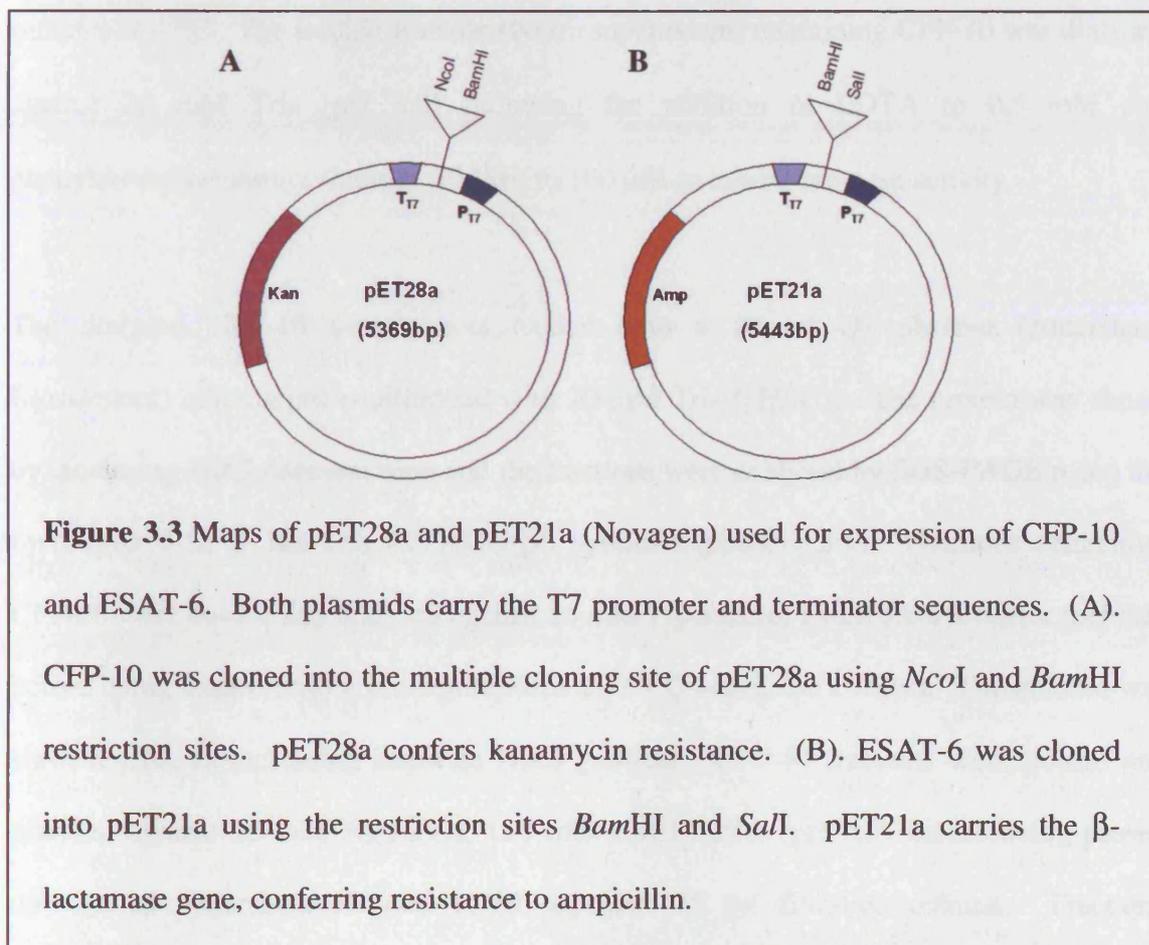


3.2 Materials and Methods

3.2.1 Protein Expression and Purification

3.2.1.1 Protein Expression Vectors

The pET28a-based expression vector for CFP-10 and the pET21a-based vector for ESAT-6 were generated by P. Renshaw (Department of Biochemistry, University of Leicester) and A. Whelan (TB Research Group, Veterinary Laboratories Agency) (Renshaw *et al.*, 2002). In summary the full length coding sequence of CFP-10 was cloned into the pET28a vector (Novagen) using the *NcoI* and *BamHI* sites and full length ESAT-6 was cloned into the pET21a vector (Novagen) using the *NdeI* and *SalI* restriction sites (Figure 3.3).



3.2.1.2 Expression and Purification of CFP-10

E. coli BL21 (DE3) (Appendix 1.1) cells transformed with the pET28a-based expression vector for CFP-10 were grown in LB media containing 40 $\mu\text{g/ml}$ kanamycin (Appendix

1.2). The cells were grown at 37 °C and 200 rpm in an orbital shaker until mid-log phase, an optical density of 0.6 at 600 nm, before induction with isopropyl-1-thio-B-D-galactoside (IPTG) to a final concentration of 0.45 mM. Cells were harvested after four hours by centrifugation at 7,800 g for 15 minutes at 4 °C. The supernatant was removed and the cell pellet resuspended in lysis buffer containing 50 mM Tris, 2 mM EDTA, 0.1 % Triton X-100 (v/v) (pH 8.0). Lysozyme was added to a final concentration of 0.1 mg/ml prior to incubation at room temperature on a rocker for 20 minutes. Samples were sonicated using an intermediate probe at full power for 30 seconds followed by 30 seconds rest. Sonication was repeated for four cycles before centrifugation at 12,100 g for 15 minutes at 4 °C. The soluble fraction (lysate supernatant) containing CFP-10 was dialysed against 20 mM Tris (pH 8.0) following the addition of EDTA to 0.5 mM and phenylmethylsulphonyl fluoride (PMSF) to 100 µM to inhibit protease activity.

The dialysed CFP-10 sample was loaded onto a 10 ml Q-sepharose (Amersham Biosciences) column pre-equilibrated with 20 mM Tris (pH 8.0). The protein was eluted by increasing NaCl concentration and the fractions were analysed by SDS-PAGE using the Invitrogen 4-12 % Bis-Tris NuPAGE gel system (Appendix 3.5). Fractions containing CFP-10 were pooled and dialysed against 20 mM Piperazine, 1 mM EDTA buffer (pH 5.8) before being loaded onto a pre-equilibrated 10 ml Q-sepharose column. The column was washed with an increasing stepwise NaCl gradient. CFP-10 fractions were pooled and dialysed against 25 mM Na₂HPO₄, 150 mM NaCl buffer (pH 6.5) before being passed through an Amersham HiLoad 16/60 superdex 75 gel filtration column. Fractions containing purified CFP-10 were pooled and the protein concentration was quantified by UV absorption at 280 nm using a Cary Bio 300 UV-visible spectrophotometer (Varian).

3.2.1.3 Expression and Purification of ESAT-6

E. coli BL21 (DE3) cells transformed with the pET21a-based expression vector for ESAT-6 were grown in LB supplemented with 100 µg/ml ampicillin (Appendix 1.2). As described for CFP-10 expression, the cells were induced with IPTG during mid-log phase and grown for a further 4 hours before being harvested by centrifugation. Cell pellets were resuspended in lysis buffer and sonicated as described for CFP-10. ESAT-6 is present in inclusion bodies in this expression system. The insoluble fraction (pellet) was collected by centrifugation at 12,100 g for 15 minutes at 4 °C. The inclusion bodies were washed and homogenised in 50 mM Tris, 10 mM EDTA, 0.5 % Triton X-100 (v/v) (pH 8.0) before being centrifuged at 12,100 g for 15 minutes at 4 °C. This process was repeated at least three times before solubilising the ESAT-6 inclusion bodies in 6 M Guanidine Hydrochloride (Gdn-HCl), 1 mM EDTA and 100 µM PMSF.

The ESAT-6 sample was refolded by dialysis against 20 mM Bis-Tris (pH 6.5) and applied to a pre-equilibrated 10 ml Q-sepharose column (Amersham Biosciences). The column was washed with an increasing stepwise gradient of NaCl and fractions were analysed by SDS-PAGE (Appendix 3.5). Pooled ESAT-6 fractions were dialysed against 20 mM Bis-Tris buffer pH 6.5 before repeating the Q-sepharose purification step. As described for CFP-10, the final polishing step involved dialysing the ESAT-6 sample against 25 mM Na₂HPO₄, 150 mM NaCl buffer (pH 6.5) followed by gel filtration. Fractions containing purified ESAT-6 were pooled and quantified by UV absorption at 280 nm.

3.2.2 Expression and Purification of MPB70

MPB70 was expressed and purified as previously described by Bloemink *et al.*, 2001. Briefly, MPB70 was purified from the periplasmic fraction of *E. coli* DH5α cells (Appendix 1.1) transformed with pBluescript KS⁺ containing the full length coding

sequence of MPB70 (pVW500, Hewinson and Russell, 1993). DH5 α cells were grown in LB supplemented with 150 μ g/ml ampicillin (Appendix 1.2) at 37 °C and 200 rpm in an orbital shaker until reaching an approximate OD of 1.7-2.0 at 600 nm. Cells were pelleted by centrifugation at 6000 rpm at 4 °C for 10 minutes. The cell pellets were resuspended in 30 mM Tris, 20 % sucrose, 1 mM EDTA (pH 8.0) and stirred at room temperature for 10 minutes. The sample was centrifuged at 6000 rpm at 4 °C for 10 minutes and the pellet resuspended in 5 mM ice-cold MgSO₄ and stirred at room temperature for 10 minutes. The sample was centrifuged as in the previous step and the supernatant containing the periplasmic fraction was retained.

The periplasmic fraction was concentrated and dialysed against 20 mM Tris (pH 8.0) before being loaded onto a 20 ml Q-sepharose (Amersham Biosciences) column pre-equilibrated with the Tris buffer and eluted with an increasing stepwise NaCl gradient. Fractions were analysed by SDS-PAGE and stained using the Silver Xpress Silver Staining Kit (Invitrogen) (Appendix 3.5), and those containing MPB70 were pooled and dialysed against 20 mM Na₂HPO₄ (pH 7.2). Prior to loading the sample onto the 1 ml Resource-Phe column (Amersham Biosciences) (NH₄)₂SO₄ was added to a final concentration of 1 M. MPB70 was eluted by a linear gradient of (NH₄)₂SO₄, from 1 M to 0 M. MPB70 fractions were dialysed against 25 mM Na₂HPO₄, 100 mM NaCl buffer (pH 6.0) before gel filtration. The protein concentration was quantified by UV absorption at 280 nm.

3.2.3 Fluorescent Labelling of the CFP-10•ESAT-6 Complex and MPB70

The CFP-10•ESAT-6 complex was prepared by mixing equimolar solutions (40 μ M) of each protein at room temperature for 30 minutes in 25 mM Na₂HPO₄, 150 mM NaCl buffer (pH 6.5). Both the complex and MPB70 were dialysed against 25 mM Na₂HPO₄, 100 mM NaCl buffer (pH 7.5) prior to labelling with Alexa Fluor 546 (Molecular Probes). The

reaction was carried out at pH 7.5 to favour reaction of the succinimidyl ester group of the Alexa Fluor dye with the N-terminal amino groups of CFP-10 and ESAT-6 or MPB70, but not the charged lysine side chain amino groups. A molar ratio of 10:1 dye to protein was used and the reaction was incubated overnight, in the dark, at room temperature with continuous rocking. Excess fluorophore was removed from the protein-dye conjugate by dialysis against 25 mM Na₂HPO₄, 100 mM NaCl buffer (pH 7.5). The degree of labelling was determined by measuring the absorbance of the protein-dye conjugate at 280 (absorbance of the protein) and 556 nm (absorbance of Alexa Fluor 546) using a Cary Bio 300 UV-visible spectrophotometer (Varian). The absorbance at A₂₈₀ was corrected for the contribution of the Alexa Fluor dye (where 0.12 is the correction factor for Alexa Fluor 546) as described by the suppliers and summarised below (Haugland *et al.*, 2005).

$$A_{\text{protein}} = A_{280} - A_{556} (0.12)$$

$$\text{Degree of labelling} = \frac{[\text{Alexa Fluor 546}]}{[\text{Protein}]}$$

3.2.4 Cell Culture

3.2.4.1 Primary Cells

Primary monocyte and macrophage cells were provided by Dr Bernard Burke (Department of Infection, Immunity and Inflammation, University of Leicester). Peripheral blood mononuclear cells were isolated from whole blood using Ficoll-Paque Plus (Amersham Biosciences) as described by Burke *et al.* (2003).

3.2.4.2 Monocytic Cell Lines

MM6 cells, provided by Dr Bernard Burke (Department of Infection, Immunity and Inflammation, University of Leicester), were cultured in RPMI 1640 media (Invitrogen)

containing 10 % foetal calf serum (FCS), 2 mM glutamine, 200 U/ml penicillin, 200 µg/ml Streptomycin, 1 X non-essential amino acids (Invitrogen), 1 X OPI supplement (Sigma). Cells were incubated at 37 °C and 5 % CO₂. Stock U937 cell cultures were kindly provided by Dr Jim Norman (Department of Biochemistry, University of Leicester) and were grown in suspension in RPMI 1640 media (Invitrogen) supplemented with 10 % FCS, 2 mM glutamine, 10 IU/ml penicillin and 10 mg/ml streptomycin. The cells were incubated at 37 °C and 5 % CO₂.

3.2.4.3 Fibroblast Cell Lines

NIH-3T3 cells were provided by Dr Patrick Caswell (Department of Biochemistry, University of Leicester) and stock Cos-1 cell cultures were provided by Dr Jim Norman (Department of Biochemistry, University of Leicester). Both cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen) supplemented with 10 % (v/v) FCS and 100 U/ml penicillin, 100 µg/ml streptomycin. Both cell lines were grown at 37 °C and 10 % CO₂.

3.2.5 Fluorescence Microscopy

Primary monocyte, monocyte-derived macrophages, NIH-3T3 and Cos-1 cells were grown directly on glass coverslips in the appropriate media. The MM6 and U937 monocytic cells were initially grown in suspension, and following removal of FCS from the media by washing the cells in RPMI minus FCS, the cells were allowed to adhere to glass coverslips pre-coated with 160 µg/ml poly-L-lysine for 20 minutes at 37 °C and 5 % CO₂. Excess cells were removed by suction and the coverslips washed twice with phosphate buffered saline (PBS). To determine whether the full length CFP-10•ESAT-6 complex binds to the surface of host cells, cells adhered to glass coverslips were exposed to 0.1 - 1 µM Alexa

Fluor 546 labelled CFP-10•ESAT-6 complex for 15 minutes at 4 °C on a rocker in the dark. Any unbound protein was removed by two washes with PBS prior to fixing the cells with 4 % (w/v) paraformaldehyde (PFA) and permeabilisation with 0.2 % (v/v) Triton X-100. Coverslips were mounted onto slides using ProLong Antifade reagent (Molecular Probes) and stored at room temperature in the dark until dry. Fluorescence microscopy was carried out using a Nikon TE300 inverted microscope and images were captured using an ORCA ER 12-bit CCD camera (1344 x 1024 pixels) (C4742-95-12ER, Hamamatsu) and Openlab software (Improvision). Fluorescence background was normalised using the Openlab software. Images were imported into Paint Shop Pro (Version 7.6) (Jasc) and were subjected to minimal processing for printing purposes.

The effects of adding an excess of unlabelled CFP-10•ESAT-6 complex were assessed with U937 cells. The cells were initially attached to glass coverslips, as described for the binding experiments above, and then incubated with a solution containing 1 µM Alexa Fluor 546 labelled CFP-10•ESAT-6 complex and a 20-fold molar excess of unlabelled CFP-10•ESAT-6 complex for 15 minutes at 4 °C. The cells were washed, fixed and permeabilised as described previously, prior to visualisation under the microscope. To determine the specificity of the interaction, MM6 cells were also exposed to 0.01-1.0 µM Alexa Fluor 546 labelled MPB70, an unrelated *M. tuberculosis* complex secreted protein.

Time course studies were used to investigate whether the CFP-10•ESAT-6 complex is internalised into the cells, and also to determine whether incubation with the complex has any noticeable effect on cell morphology or viability. MM6 cells were incubated with the complex for 30 minutes at 4 °C on a rocker in the dark. Unbound protein was removed by washing the cells with PBS before incubating the cells in RPMI media for 15, 30 or 60 minutes at 37 °C and 5 % CO₂. Following the incubation, the media was removed and the

cells washed in PBS. Prior to viewing under the microscope, the cells were fixed and permeabilised.

3.3 Results

3.3.1 Fluorescent Labelling of the CFP-10•ESAT-6 Complex and MPB70

3.3.1.1 Expression and Purification of CFP-10

Using the pET28a-based system CFP-10 is produced as a soluble protein, and can be found in the supernatant of the whole cell lysate (soluble fraction) (Figure 3.4).

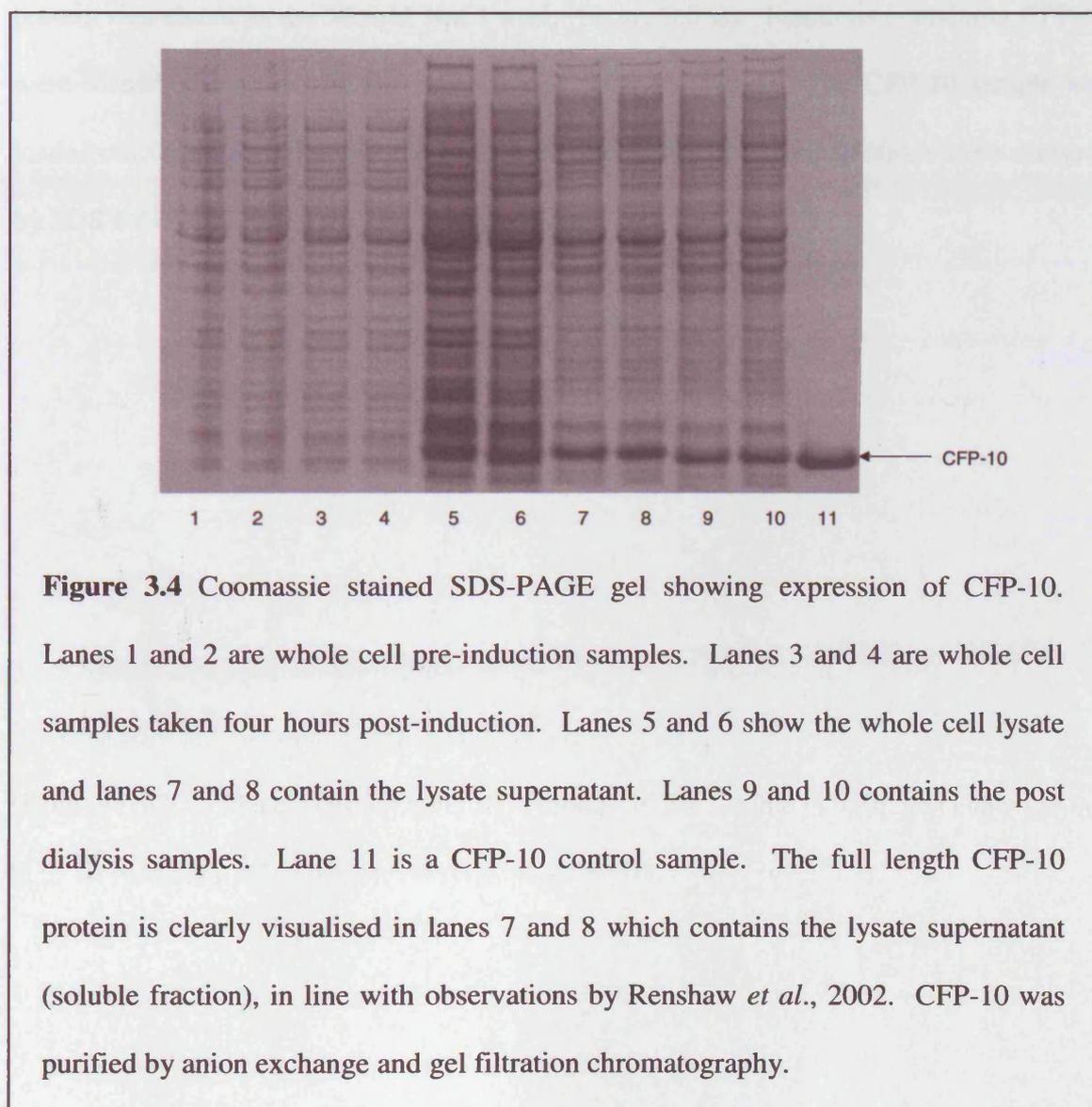


Figure 3.4 Coomassie stained SDS-PAGE gel showing expression of CFP-10.

Lanes 1 and 2 are whole cell pre-induction samples. Lanes 3 and 4 are whole cell samples taken four hours post-induction. Lanes 5 and 6 show the whole cell lysate and lanes 7 and 8 contain the lysate supernatant. Lanes 9 and 10 contains the post dialysis samples. Lane 11 is a CFP-10 control sample. The full length CFP-10 protein is clearly visualised in lanes 7 and 8 which contains the lysate supernatant (soluble fraction), in line with observations by Renshaw *et al.*, 2002. CFP-10 was purified by anion exchange and gel filtration chromatography.

CFP-10 was initially purified by anion exchange chromatography. The whole cell lysate containing the CFP-10 protein was dialysed against 20 mM Tris, 1 mM EDTA (pH 8.0) buffer before being loaded onto a Q-sepharose anion exchange column. The column was washed with an increasing stepwise NaCl gradient and CFP-10 was eluted in the 75-100 mM NaCl wash (Figure 3.5 A). Based on SDS-PAGE analysis, samples containing CFP-10 were pooled and dialysed against 20 mM Piperazine, 1 mM EDTA buffer (pH 5.8) before being reloaded onto a Q-sepharose column pre-equilibrated with the base buffer. The column washed with an increasing stepwise gradient of NaCl. In this step the CFP-10 protein was eluted in the 50 mM NaCl wash (Figure 3.5 B). Fractions containing CFP-10 were identified by SDS-PAGE analysis and pooled together. The CFP-10 sample was loaded onto a gel filtration column as a final polishing step. Eluted fractions were analysed by SDS-PAGE (Figure 3.6).

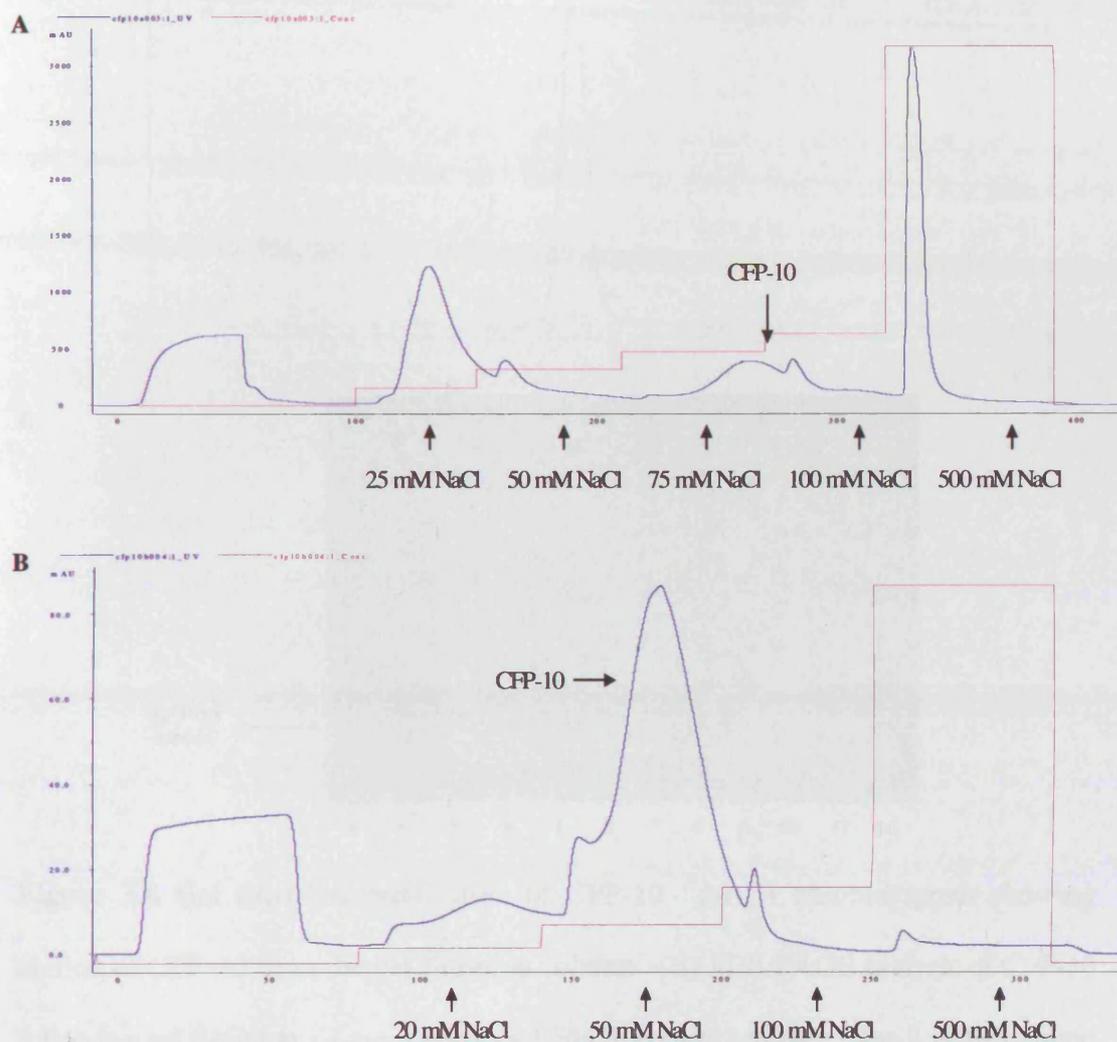
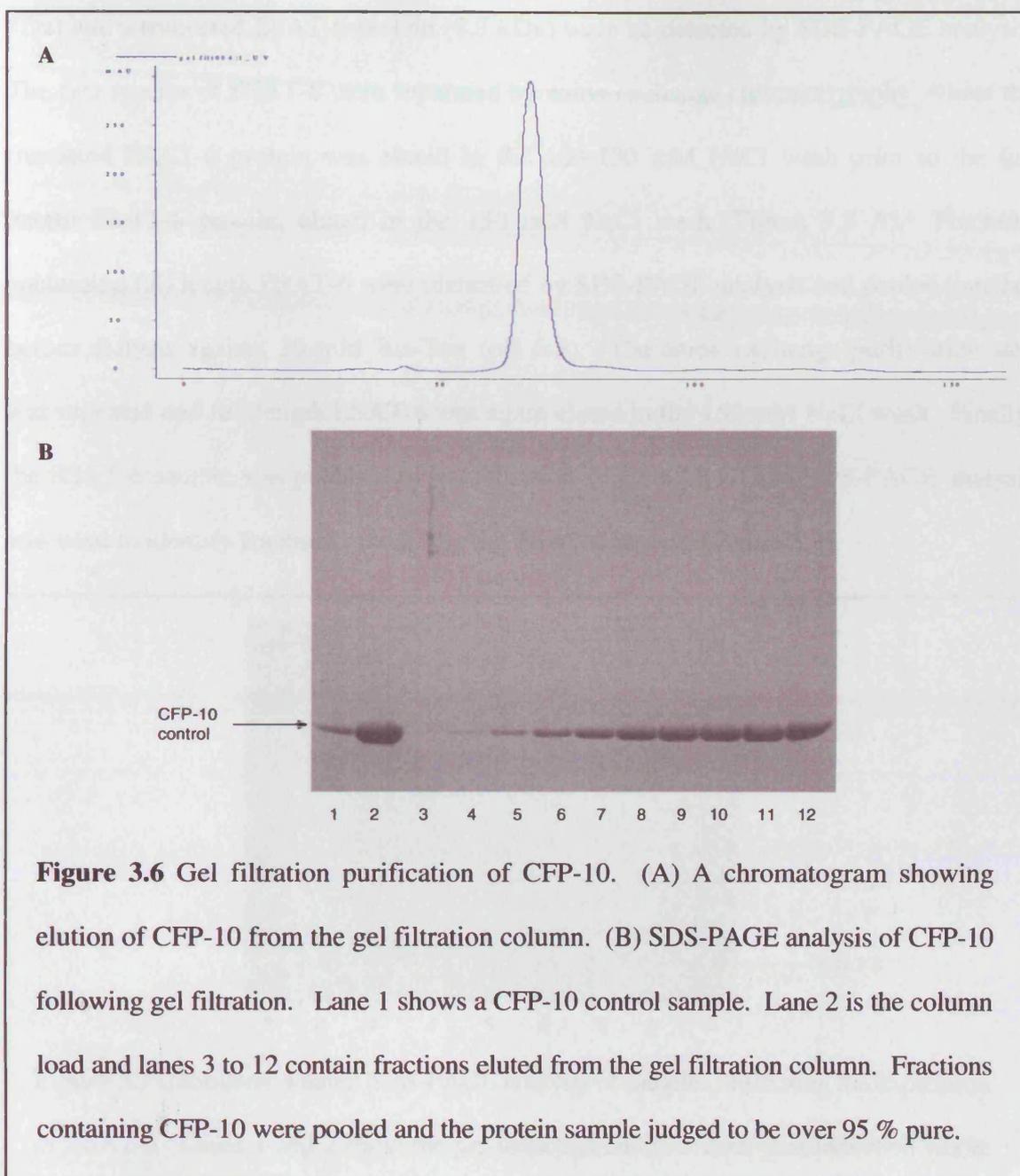


Figure 3.5 Anion exchange purification elution profiles for CFP-10. (A) The base buffer for the initial purification step was 20 mM Tris, 1 mM EDTA (pH 8.0). In this step CFP-10 was eluted in the 75-100 mM NaCl wash. (B) In the second purification step CFP-10 was predominantly eluted in the 50 mM NaCl wash. The base buffer for this step was 20 mM Piperazine, 1 mM EDTA (pH 5.8)



3.3.1.2 Expression and Purification of ESAT-6

In contrast to CFP-10, ESAT-6 was produced as insoluble inclusion bodies (Figure 3.7). The inclusion bodies containing ESAT-6 were solubilised in buffer containing Gdn-HCl, and the protein was subsequently refolded by removal of Gdn-HCl by dialysis against 20 mM Bis-Tris (pH 6.5). As described for CFP-10, ESAT-6 was initially purified by anion exchange chromatography. Following dialysis, both the full length ESAT-6 protein (9.9

kDa) and a truncated ESAT-6 protein (8.8 kDa) were detected by SDS-PAGE analysis. The two species of ESAT-6 were separated by anion exchange chromatography, where the truncated ESAT-6 protein was eluted in the 100-150 mM NaCl wash prior to the full length ESAT-6 protein, eluted in the 150 mM NaCl wash (Figure 3.8 A). Fractions containing full length ESAT-6 were identified by SDS-PAGE analysis and pooled together before dialysis against 20 mM Bis-Tris (pH 6.5). The anion exchange purification step was repeated and full length ESAT-6 was again eluted in the 150 mM NaCl wash. Finally, the ESAT-6 sample was polished by gel filtration (Figure 3.8 B) and SDS-PAGE analysis was used to identify fractions containing the ESAT-6 protein (Figure 3.9).

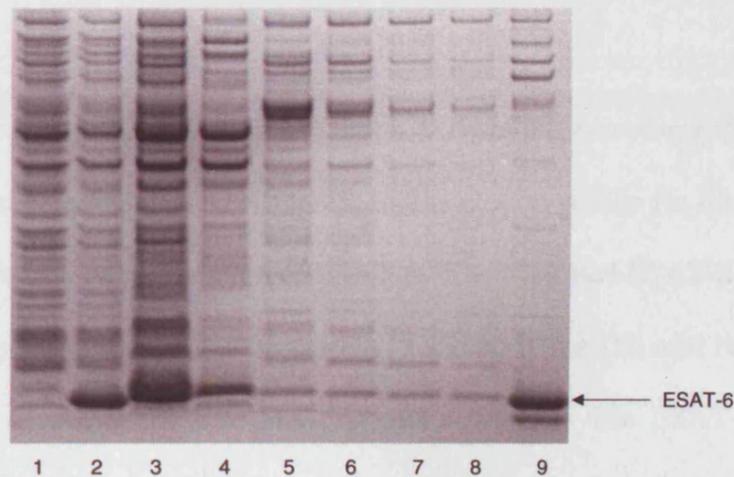
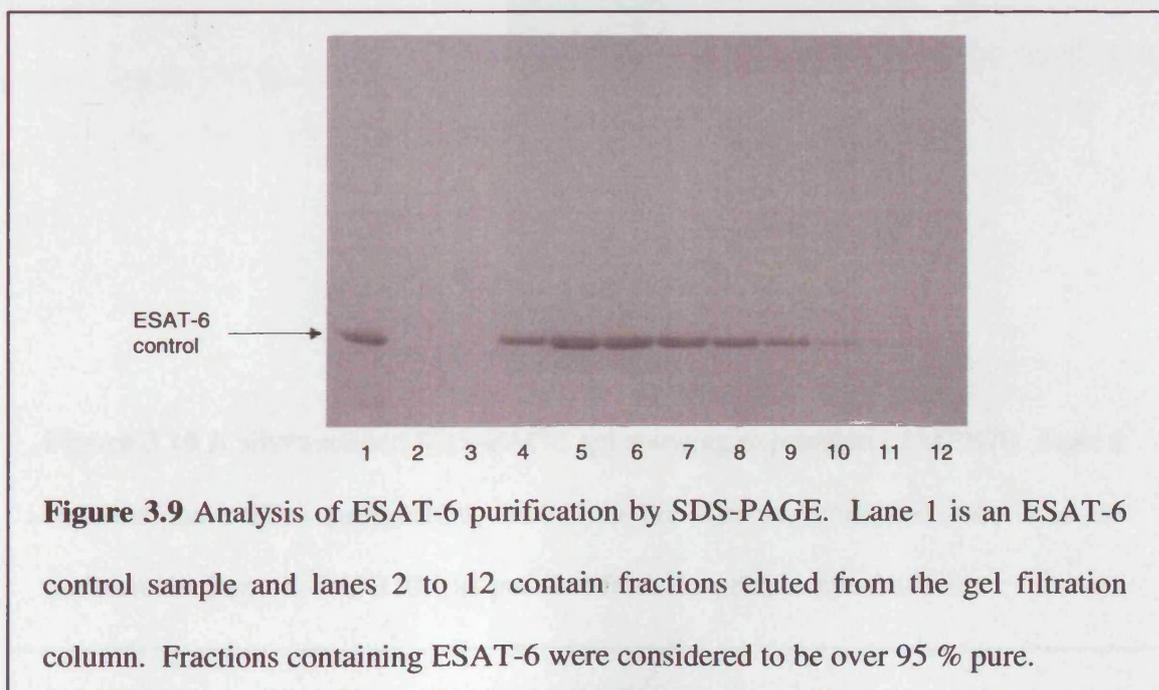
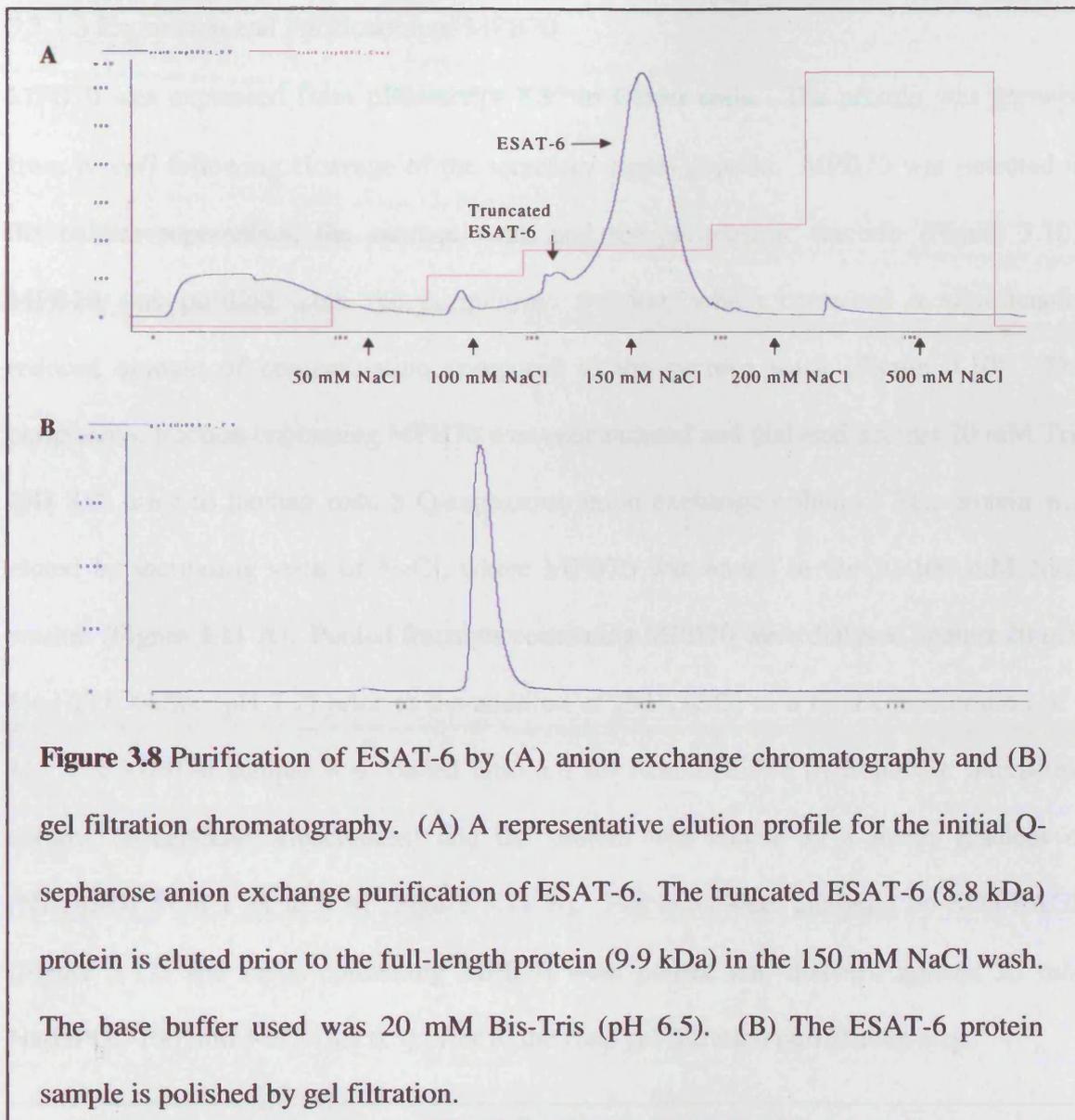


Figure 3.7 Coomassie stained SDS-PAGE analysis of samples illustrating the expression of ESAT-6. Lanes 1 and 2 show the pre-induction and four hour post-induction whole cell samples. Lanes 3 and 4 contain the total cell lysate and lysate supernatant. Lanes 5, 6, 7 and 8 contain samples from the inclusion body washes. Lane 9 is the post-dialysis sample. ESAT-6 is clearly present in the total cell lysate, where it is found as insoluble inclusion bodies. The inclusion body washes in lanes 5 to 8 clearly show the presence of the full length ESAT-6 protein (9.9 kDa). Truncated ESAT-6 (8.8 kDa) can be detected along with the full length ESAT-6 protein (9.9 kDa) in the post dialysis sample in lane 9. The full length ESAT-6 protein was purified by anion exchange and gel filtration to remove contaminants, including the truncated ESAT-6 species.



3.3.1.3 Expression and Purification of MPB70

MPB70 was expressed from pBluescript KS⁺ in DH5 α cells. The protein was secreted from *E. coli* following cleavage of the secretory signal peptide. MPB70 was detected in the culture supernatant, the sucrose wash and the periplasmic fraction (Figure 3.10). MPB70 was purified from the periplasmic fraction, which contained a significantly reduced amount of contamination compared to the sucrose wash (Figure 3.10). The periplasmic fraction containing MPB70 was concentrated and dialysed against 20 mM Tris (pH 8.0) prior to loading onto a Q-sepharose anion exchange column. The protein was eluted by increasing steps of NaCl, where MPB70 was eluted in the 50-100 mM NaCl washes (Figure 3.11 A). Pooled fractions containing MPB70 were dialysed against 20 mM Na₂HPO₄ buffer (pH 7.2) prior to the addition of (NH₄)₂SO₄ to a final concentration of 1 M. The MPB70 sample was loaded onto a 1 ml Resource-Phe hydrophobic interaction column (Amersham Biosciences) and the protein was eluted by a linear gradient of (NH₄)₂SO₄ from 1 M to 0 M (Figure 3.11 B). Fractions were analysed by SDS-PAGE (Figure 3.12) and those containing MPB70 were pooled and dialysed against 25 mM Na₂HPO₄, 100 mM NaCl (pH 6.5) prior to the final gel filtration purification step.

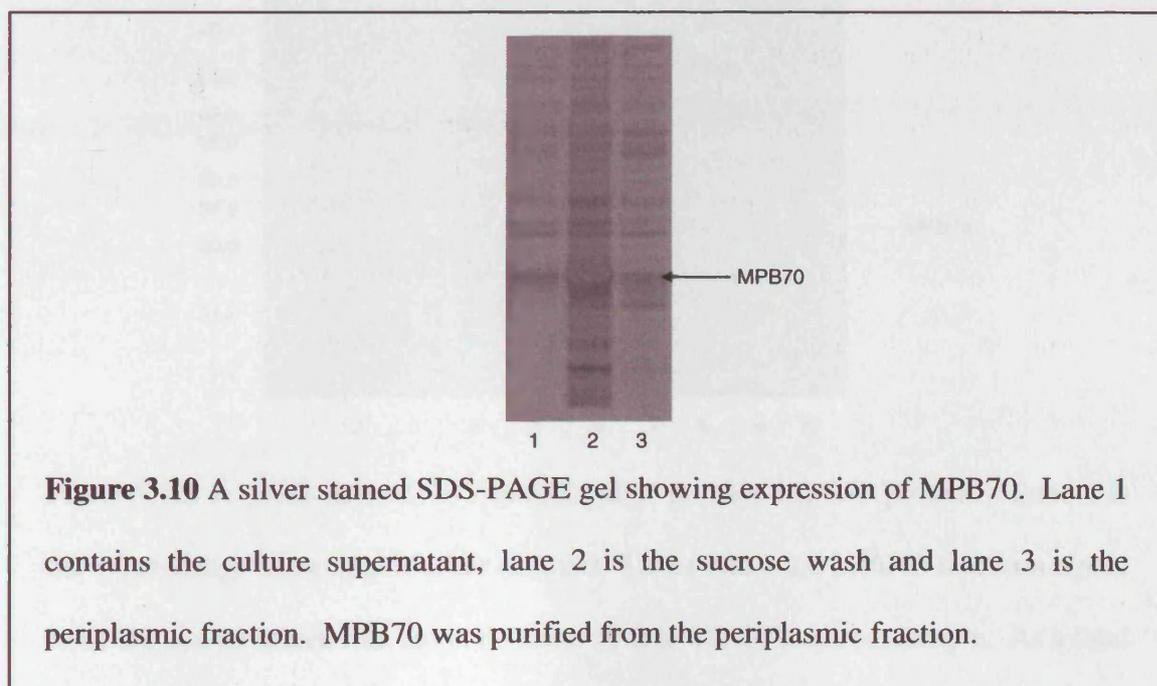


Figure 3.10 A silver stained SDS-PAGE gel showing expression of MPB70. Lane 1 contains the culture supernatant, lane 2 is the sucrose wash and lane 3 is the periplasmic fraction. MPB70 was purified from the periplasmic fraction.

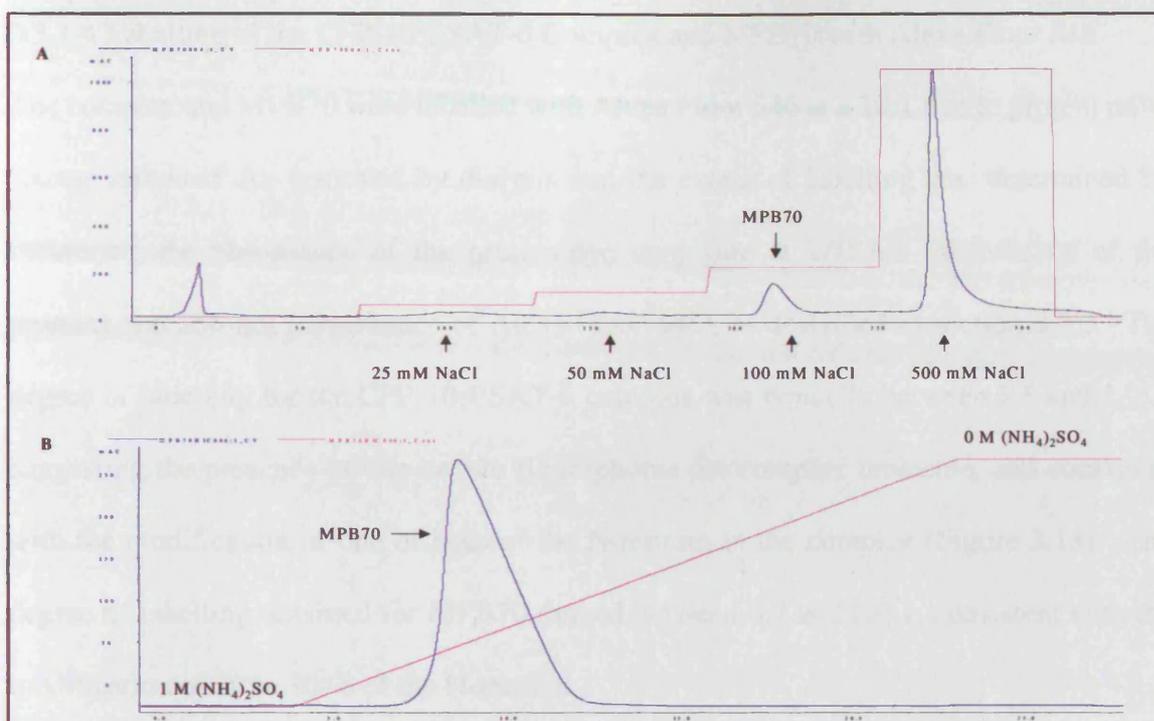


Figure 3.11 Anion exchange purification of MPB70. (A) Initial purification of MPB70 by anion exchange chromatography, using a Q-sepharose column and 20 mM Tris base buffer (pH 8.0). MPB70 is eluted in the 50-100 mM NaCl wash. (B) MPB70 is eluted from the Resource-Phe hydrophobic interaction column by a linear gradient of $(\text{NH}_4)_2\text{SO}_4$ from 1 M to 0 M.

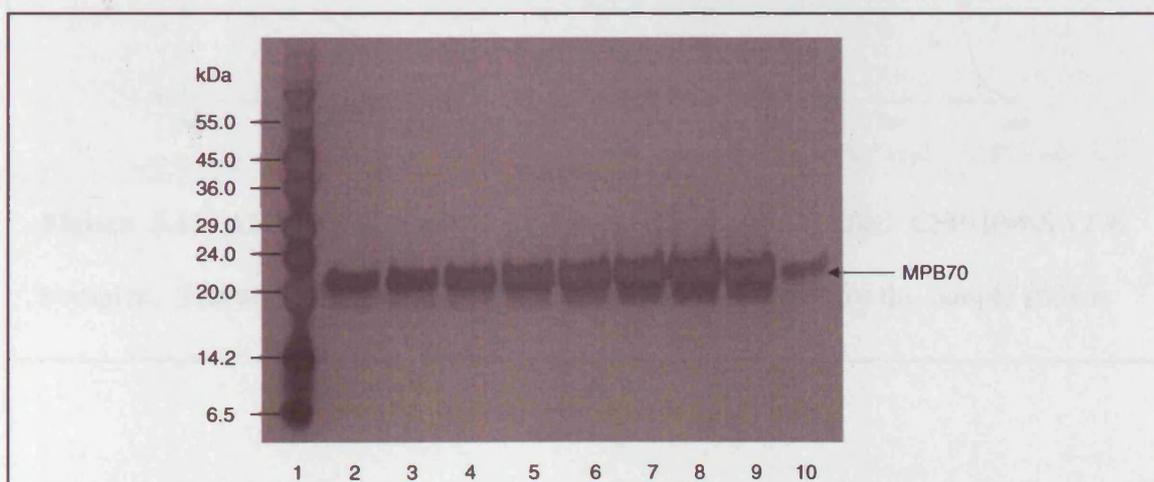
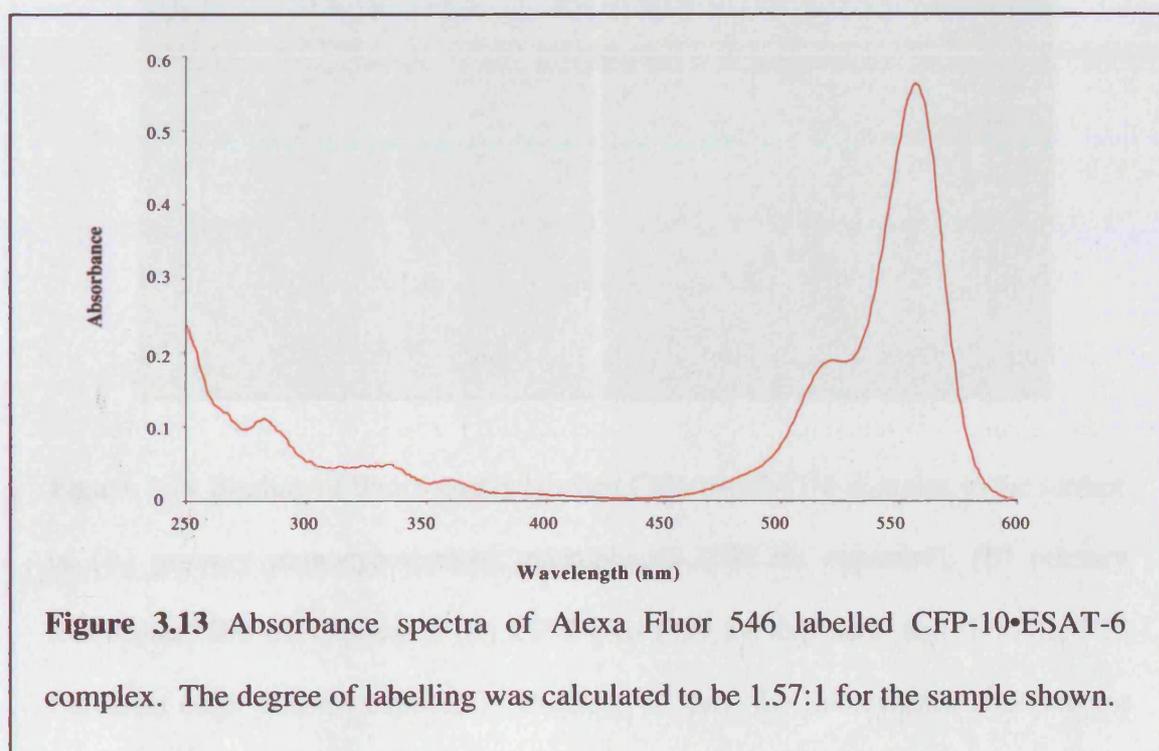


Figure 3.12 A silver stained SDS-PAGE gel of purified MPB70 protein. Lane 1 is the Wide Range Molecular Marker (Sigma). Lanes 2 to 9 are MPB70 fractions eluted from the 1ml Resource-Phe column. Lane 10 is an MPB70 control sample. As a final step the pooled fractions containing MPB70 were polished by gel filtration.

3.3.1.4 Labelling of the CFP-10•ESAT-6 Complex and MPB70 with Alexa Fluor 546

The complex and MPB70 were labelled with Alexa Fluor 546 at a 10:1 dye to protein ratio. Excess unbound dye removed by dialysis and the extent of labelling was determined by measuring the absorbance of the protein-dye conjugate at 280 nm (absorbance of the protein) and 556 nm (absorbance of Alexa Fluor 546), as described in section 3.2.3. The degree of labelling for the CFP-10•ESAT-6 complex was typically between 1.5 and 1.9:1, suggesting the presence of one or two fluorophores per complex molecule, and consistent with the modification of one or both of the N-termini in the complex (Figure 3.13). The degree of labelling obtained for MPB70 ranged between 0.7 and 0.9:1, consistent with the modification of 70 to 90 % of the N-termini.



3.3.2 Fluorescence Microscopy

3.3.2.1 Binding of the CFP-10•ESAT-6 Complex to Specific Host Cells

Fluorescently labelled CFP-10•ESAT-6 complex was used to determine whether the complex bound specifically to a variety of host cells including, primary human monocyte

and macrophage cells, monocytic cell lines (U937 and MM6) and fibroblast cell lines (NIH-3T3 and COS-1). The primary monocyte/macrophage cells together with both monocytic cell lines consistently showed intense fluorescence at the cell surface, indicating binding of the CFP-10•ESAT-6 complex (Figure 3.14 A-C). In contrast, no significant fluorescent labelling was observed with either fibroblast cell line (Figure 3.14 D).

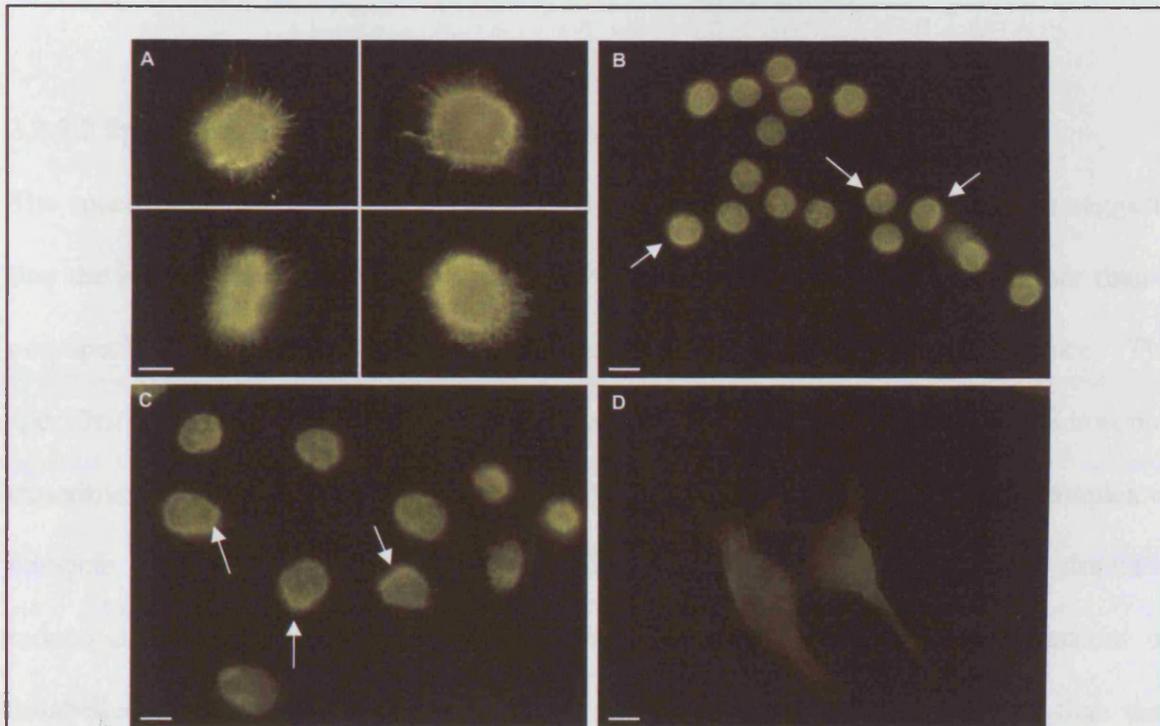


Figure 3.14 Binding of fluorescently labelled CFP-10•ESAT-6 complex to the surface of (A) primary monocyte-derived macrophages (100 ms exposure), (B) primary monocytes (200 ms exposure), (C) U937 cells (100 ms exposure) and (D) NIH-3T3 fibroblast cells (200 ms exposure). Panels A-C show the fluorescence observed for monocyte lineage cells (A-C) exposed to 1 μ M Alexa Fluor 546 labelled CFP-10•ESAT-6. Labelling is focussed at the cell surface and is often observed in patches (indicated by arrows), similar to the 'cap-like' structures associated with receptor mediated signalling. Typically, fibroblast cells, represented in panel (D) by NIH-3T3 cells, showed no significant labelling at the cell surface. The size bars shown correspond to 5 μ m.

Fluorescence labelling of the monocyte lineage cells was frequently focussed in 'cap-like' structures, which are often associated with cell surface receptors (Junge *et al.*, 1999, Kwiatkowska and Sobota, 1999b). The results obtained suggest that the Alexa Fluor 546 labelled complex binds to specific host cell types, implying the presence of a specific receptor for CFP-10•ESAT-6 on the surface of monocytic lineage cells but not on fibroblast cell types.

3.3.2.2 Specificity of CFP-10•ESAT-6 Binding to Host Cells

The specific binding of the CFP-10•ESAT-6 complex to monocyte lineage cells suggests that the interaction observed is the result of binding to a host cell receptor rather than a non-specific interaction between the complex or fluorophore and the cell surface. The specificity of the interaction was confirmed by comparable fluorescence microscopy experiments carried out in the presence of a 20-fold molar excess of unlabelled complex to compete with Alexa Fluor 546 labelled CFP-10•ESAT-6. This revealed a dramatic reduction in cell surface associated fluorescence for cells exposed to an excess of unlabelled complex, which appeared to be several orders of magnitude lower than that observed for cells exposed to labelled CFP-10•ESAT-6 complex alone (Figure 3.15). This clearly indicates that binding to the cell surface is mediated by the protein complex and not by the fluorophore. Further experiments where primary monocytes/macrophages and MM6 cells were exposed to Alexa Fluor 546 labelled MPB70 showed no significant labelling at the cell surface, as shown in figure 3.16, confirming that the labelled CFP-10•ESAT-6 complex binds to a specific receptor on the surface of monocyte cells, and that the binding is not mediated by the Alexa Fluor dye.

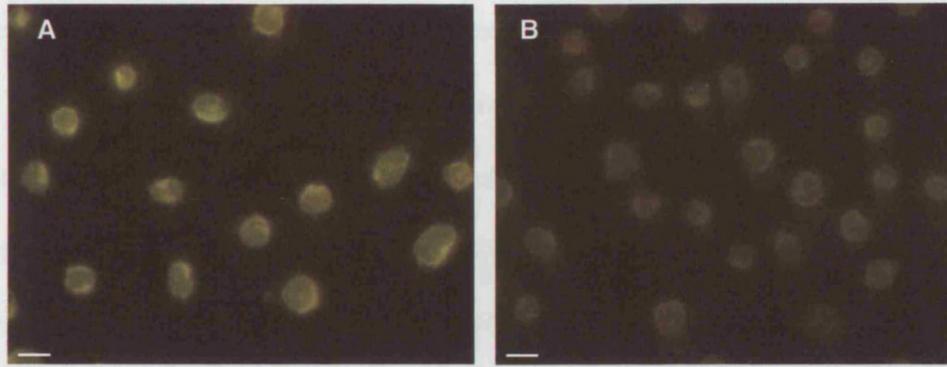


Figure 3.15 Inhibiting the binding of fluorescently labelled CFP-10•ESAT-6 to U937 cells by competition with unlabelled complex. (A) Typical labelling of U937 cells with 1 μ M Alexa Fluor 546 labelled CFP-10•ESAT-6. (B) U937 cells exposed to 1 μ M labelled CFP-10•ESAT-6 and 20 μ M unlabelled complex. It is clear that fluorescence labelling in the presence of unlabelled CFP-10•ESAT-6 complex is drastically reduced in comparison to labelled complex alone, indicating that the interaction between the complex and the host cell surface is not fluorophore mediated. The size bars shown correspond to 10 μ m.

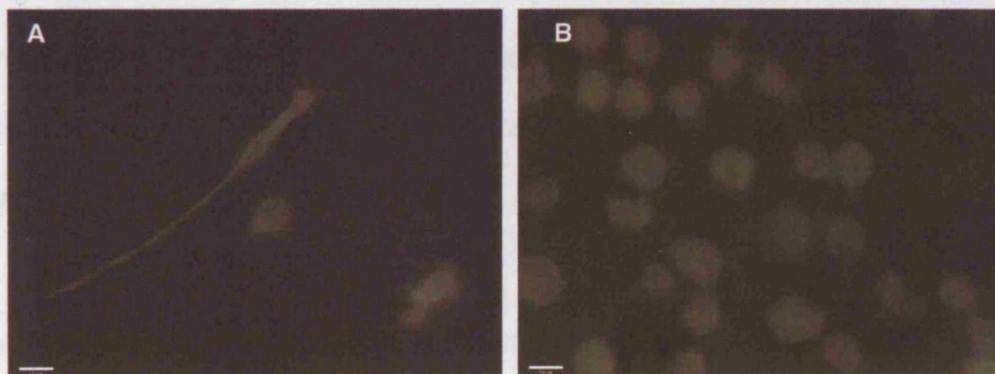


Figure 3.16 Interaction of Alexa Fluor 546 labelled MPB70 with (A) primary monocytes/macrophages and (B) MM6 monocyte cells. No significant fluorescent labelling was observed at the cell surface, indicating that MPB70 does not bind to the surface of monocytic lineage cells. The size bars shown correspond to 10 μ m

3.4 Discussion

The microscopy studies described in this thesis provide the first evidence that the CFP-10•ESAT-6 complex specifically binds to a receptor present on the surface of monocyte-lineage cells. A previous study indicated that the individual CFP-10 protein binds to the surface of monocytic cells (Trajkovic *et al.*, 2002), however due to the random coil nature of the individual protein this could be the result of non-specific interactions at the cell surface. In the host cell binding assays described in this chapter a significant difference in fluorescent labelling of the surface of monocyte/macrophage lineage cells was observed in comparison to fibroblast-type cells. U937 and MM6 monocyte cell lines showed consistent labelling at the cell surface; however it was the primary monocytes and macrophages which were labelled most intensely (Figure 3.14 A-C). In contrast, there was no significant labelling of the fibroblast cell lines NIH-3T3 or Cos-1, which suggests a high degree of interaction between the CFP-10•ESAT-6 complex and monocyte and macrophage cells. Recent studies by Brodin *et al.* (2006) demonstrated that *M. tuberculosis* expressing the RD1 region were engulfed by macrophages more efficiently than *M. tuberculosis* strains lacking the RD1 region. However, the same effect was not witnessed with epithelial cells (A549) suggesting that the RD1 effectors, CFP-10 and ESAT-6, preferentially interact with monocyte/macrophage cells, which supports the interaction specificity reported here.

The obvious difference in fluorescence intensity between the various cell types suggests that the interaction between the labelled complex and monocyte cells is the result of tight binding to a specific receptor and is not due to non-specific interactions of the fluorophore, which would be similar across all cell types. The specificity of the interaction was confirmed by blocking experiments, where U937 cells were incubated with fluorescently labelled complex in the presence of a 20-fold molar excess of unlabelled CFP-10•ESAT-6.

Microscopy results demonstrated a dramatic reduction in fluorescence intensity, demonstrating that incubation with excess unlabelled complex prevented the binding of the labelled complex (Figure 3.15).

M. tuberculosis and *M. bovis* secreted immunodominant antigen MPT70/MPB70 was labelled with Alexa Fluor 546 and was also used in binding assays with MM6 and primary monocyte/macrophage cells. In these studies no significant labelling was detected on monocyte cells, suggesting that MPB70 does not bind to the cell surface. These results demonstrate that not all Alexa Fluor 546 labelled proteins bind to the surface of monocytes and confirm that the labelling observed with the Alexa Fluor 546 labelled CFP-10•ESAT-6 was not mediated by the fluorophore. In addition, these results are in line with the observations of Schnappinger *et al.* (2003) who reported that MPT70, in contrast to CFP-10 and ESAT-6, is up-regulated upon bacterial internalisation into the phagosome, which strongly suggests that MPB70 is likely to function within the host cell, and is unlikely to bind to a specific receptor on the cell surface.

In a significant proportion of monocyte cells, the fluorescence labelling was concentrated in patches resembling the 'cap-like' structures associated with cell surface receptors (Figures 3.14 B and C). Patching and capping can trigger a number of cell responses, including the induction of cell signalling pathways, particularly protein tyrosine phosphorylation, increasing intracellular Ca²⁺ concentration and phagocytosis (Liao *et al.*, 1992, Vossebeld *et al.*, 1995, Kwiatkowska and Sobota, 1999a, Kwiatkowska and Sobota, 1999b). Capping of receptors on T and B-cells has also been associated with cell activation and proliferation (Snapper *et al.*, 1991, Khan *et al.*, 1992, Junge *et al.*, 1999). The formation of 'cap-like' structures on cells exposed to the CFP-10•ESAT-6 complex strongly suggests tight binding of the complex to a specific receptor on the cell surface.

This capping effect may also be accompanied by a signalling cascade, which fits with the suggestion that the CFP-10•ESAT-6 complex functions as a pathogen to host signalling molecule.

The structural and surface features of the CFP-10•ESAT-6 complex suggest a function based on the binding to one or more target proteins (Renshaw *et al.*, 2005), and is inconsistent with the proposal that the CFP-10•ESAT-6 complex has host cell lysis activity (Hsu *et al.*, 2003). The complex is stable to over 2 mM in aqueous solution with no sign of aggregation, which is not typical of a pore forming protein. Also, time course studies exposed MM6 cells to labelled CFP-10•ESAT-6 complex for periods up to 24 hours with no evidence of cell lysis. Therefore it is clear from these studies that the CFP-10•ESAT-6 complex is not associated with cytolytic activity.

In the fluorescence microscopy assays it was observed that a significant number of primary cells exposed to the labelled CFP-10•ESAT-6 complex became detached from the glass coverslips after multiple washes with PBS. However, in preliminary experiments where primary monocyte and macrophage cells were exposed to Alexa Fluor 546 labelled Rv0287•Rv0288 complex the cells remained attached to the glass coverslip. Macrophage cells exposed to the complex for 24 hours appeared more spread-out and showed increased lamellopodia formation, compared to cells exposed to a buffer control. The spreading and formation of lamellopodia suggests that signalling via the complex could trigger cell migration, which would also explain why cells initially adhered to glass coverslips were removed by washing with PBS after exposure to the complex. This suggests that the CFP-10•ESAT-6 complex may function as a host cell signalling molecule which may act as a chemotactic factor to attract host cells to the mycobacteria or to mycobacteria infected cells. In line with this proposal, recent studies have suggested that CFP-10 and ESAT-6

may be involved in host cell signalling and modulation of host cell activity. Volkman *et al.* (2004) reported that secreted RD1 virulence factors are required for the aggregation of macrophages and the subsequent formation of granulomas in zebrafish infected with *M. marinum*, which allows increased bacterial spread between host cells.

Chapter 4

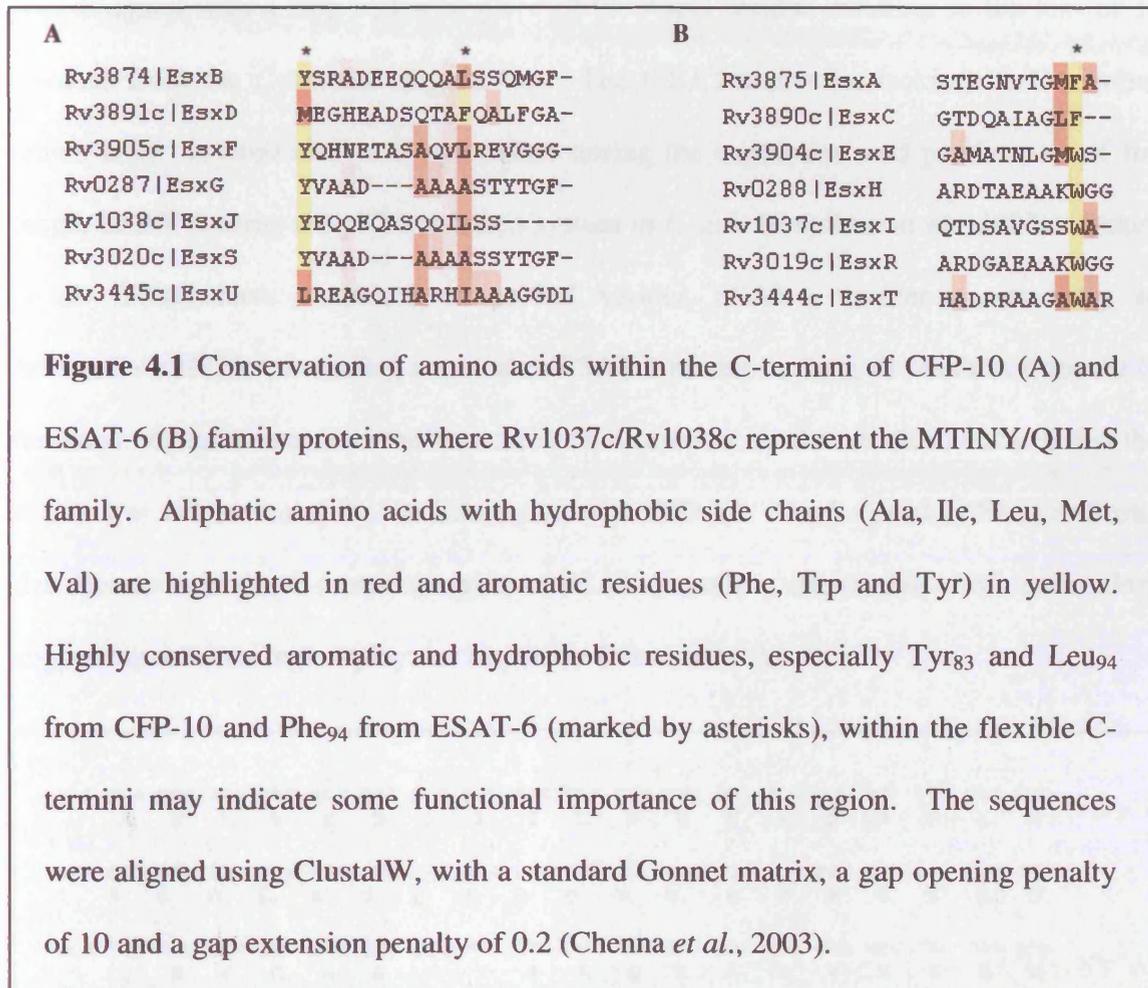
Assessment of the Functional Role of the Flexible C-termini of CFP-10 and ESAT-6

4.1 Introduction

The fluorescence microscopy experiments, described in the previous chapter, clearly demonstrate that the CFP-10•ESAT-6 complex binds specifically to the surface of monocyte lineage cells. Although the precise function of the complex remains unknown, the microscopy results, along with the structural features of the complex, strongly suggest the presence of a specific receptor, or receptors, on the surface of host cells, and a possible role in host cell signalling. Active transport of the complex by a specialised secretion system (Brodin *et al.*, 2006) and the reported down-regulation of both CFP-10 and ESAT-6 in mycobacteria within the macrophage phagosome (Schnappinger *et al.*, 2003) support the suggestion that the CFP-10•ESAT-6 complex functions outside of the host cell prior to uptake into the phagosome. Also, studies by Volkman *et al.* (2004) demonstrate that deletion of the RD1 region from *M. marinum* results in decreased granuloma formation in the zebrafish model, suggesting that secreted RD1 effector proteins (CFP-10 and ESAT-6) may be involved in macrophage aggregation and the formation of granulomas, which agrees with the proposal that CFP-10 and ESAT-6 may play a key role in pathogen-host signalling.

One of the most striking features of the CFP-10•ESAT-6 complex are the disordered N- and particularly C-termini of both proteins, which form long flexible arms at either end of the four helix bundle (Renshaw *et al.*, 2005). Sequence comparisons identified conserved hydrophobic and aromatic residues within the C-termini of CFP-10 (Tyr₈₃ and Leu₉₄) and ESAT-6 (Phe₉₄). These residues are not involved in a structural role and their presence at

the protein surface may indicate some functional significance for the C-terminal regions. Multiple sequence alignments highlighting the presence of conserved hydrophobic and aromatic residues in the C-termini of *M. tuberculosis* CFP-10/ESAT-6 family proteins are shown in figure 4.1.



The work described in this chapter aimed to determine whether the C-terminal arms of CFP-10 or ESAT-6 form an essential part of the interaction site involved in binding to a receptor, or receptors, on the surface of host cells. To test this hypothesis, complexes formed with C-terminally truncated CFP-10 (tCFP-10, residues 1-86) and ESAT-6 (tESAT-6, residues 1-84) proteins were labelled with Alexa Fluor 546 (Molecular Probes) and exposed to U937 monocytic cells. The degree of fluorescent labelling was analysed by fluorescence spectroscopy. Differences in the degree of labelling between the CFP-

10•ESAT-6, tCFP-10•ESAT-6 and CFP-10•tESAT-6 complexes should indicate whether the C-termini of CFP-10 or ESAT-6 form part of the binding site for the host cell receptor.

Based on analysis of the structure of the CFP-10•ESAT-6 complex, the tCFP-10 construct was designed with a stop codon in place of the Asp₈₇ residue resulting in the loss of 14 residues from the C-termini (Figure 4.2). The tESAT-6 protein, lacking 11 C-terminal amino acids, is produced as a by-product during the expression and purification of full length ESAT-6 using the pET21a-based system in *E. coli* (Renshaw *et al.*, 2002). Okkels *et al.* (2004) have previously identified various ESAT-6 species secreted by *M. tuberculosis* H37Rv, including a truncated ESAT-6 protein lacking 11 residues from the C-terminus. In agreement with the structure of the complex, further studies indicated that this region was not essential for the interaction with CFP-10. Brodin *et al.*, (2005) indicated that mutations in the C-terminal region of ESAT-6 result in attenuation of *M. tuberculosis* suggesting that this region plays an important functional role.

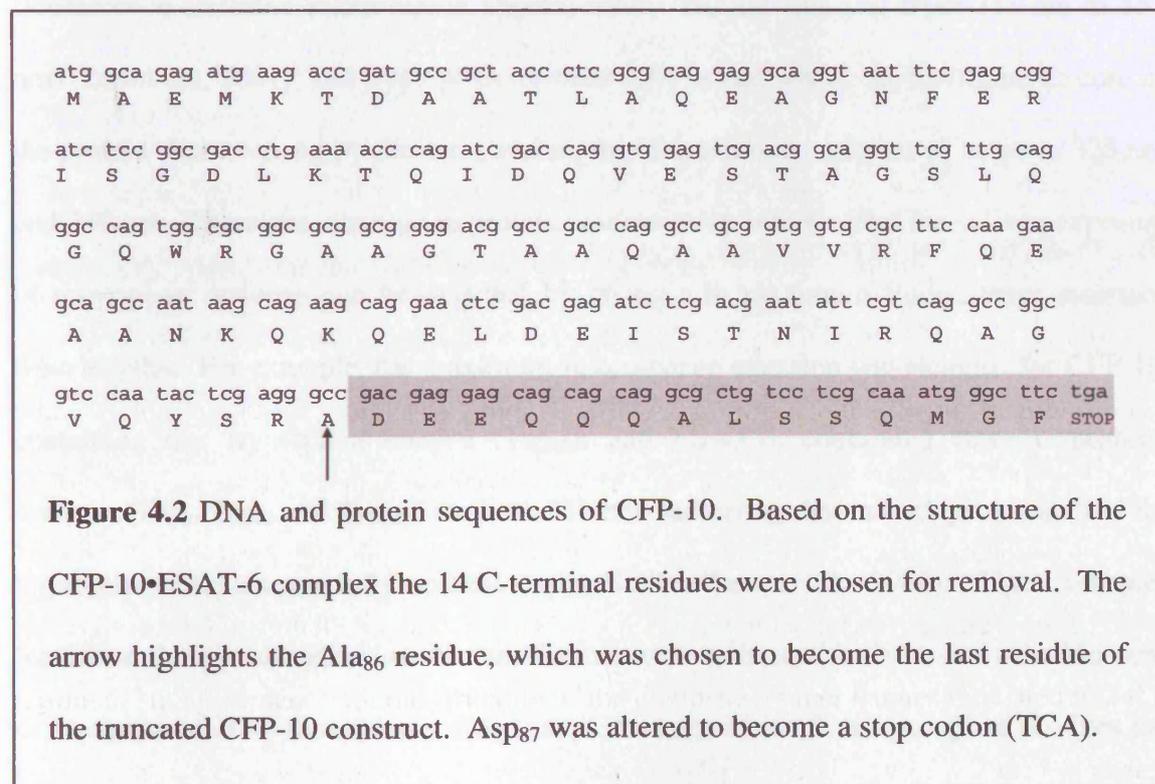


Figure 4.2 DNA and protein sequences of CFP-10. Based on the structure of the CFP-10•ESAT-6 complex the 14 C-terminal residues were chosen for removal. The arrow highlights the Ala₈₆ residue, which was chosen to become the last residue of the truncated CFP-10 construct. Asp₈₇ was altered to become a stop codon (TCA).

CFP-10 and ESAT-6 are known to form a tight 1:1 complex (Renshaw *et al.*, 2002), however, analysis of complex formation between full length and truncated versions of both CFP-10 and ESAT-6 (tCFP-10/ESAT-6, CFP-10/tESAT-6 and tCFP-10/tESAT-6) was required to determine the suitability of the truncated complexes for use in fluorescence microscopy assays. As described for CFP-10 and ESAT-6 (Renshaw *et al.*, 2002), complex formation was investigated by intrinsic tryptophan fluorescence.

The aromatic amino acids tryptophan (Trp) and tyrosine (Tyr) act as natural fluorophores allowing intrinsic protein fluorescence to be used to study protein folding and complex formation. Protein fluorescence results from excitation at 280 nm and is dominated by tryptophan emission, although tyrosine may also contribute at this wavelength (Lakowicz, 1983). The wavelength of maximum fluorescence for tryptophan residues is highly sensitive to the local environment and depends on the degree of solvent exposure; with tryptophan residues in a fully aqueous solvent exposed environment characterised by a fluorescence emission maximum at approximately 350 nm (ranging from 348 nm to 353 nm) (Ladokhin, 2001), and tryptophan residues fully buried within the hydrophobic core of the protein characterised by shorter wavelengths of maximum emission of between 325 nm and 330 nm. Therefore, changes in protein conformation which affect the solvent exposure of tryptophan residues can be detected by changes in maximum fluorescence emission wavelengths. For example, the maximum fluorescence emission wavelengths for CFP-10, containing one tryptophan residue (Trp₄₃), and ESAT-6, containing three tryptophan residues (Trp₆, Trp₄₃ and Trp₅₈), are both 353 nm, indicating that in both proteins all of the tryptophan residues are fully solvent exposed (Renshaw *et al.*, 2002). Upon complex formation the wavelength of maximum fluorescence emission shifts towards the blue end of the spectrum, to 342 nm, indicating that at least one of the four tryptophan residues has moved to a less polar environment.

4.2 Materials and Methods

4.2.1 Cloning of C-terminal Truncated CFP-10

The C-terminal tCFP-10 construct (residues 1-86) was amplified from the BAC Rv414 template (Brosch *et al.*, 1998). Forward primers were designed with a *NcoI* restriction site and reverse primers with a *BamHI* restriction site (Appendix 2). PCR reactions were carried out in 50 µl reactions using *Pfu* polymerase (Promega) as described for CFP-10 and ESAT-6 (Section 2.2.2). The tCFP-10 PCR products were analysed by gel electrophoresis using 1 % (w/v) agarose gels (Appendix 3.1) and purified using the PCR Clean-up Kit (Qiagen).

The pET28a vector (Novagen) (Figure 3.3A) and the purified PCR product were double digested with *NcoI* and *BamHI* restriction enzymes (Promega) for 7 hours at 37 °C (Appendix 3.2). Ligation reactions were carried out as described in section 2.2.2. *E. coli* DH5α (Appendix 1.1) cells successfully transformed with the ligase mix were grown on LB media supplemented with 40 µg/ml kanamycin (Appendix 1.2). Plasmid DNA from successful transformations was digested with *NcoI* and *BamHI* (as above) before analysis by electrophoresis to check for the presence of the inserted PCR product. The integrity of the inserted PCR fragment was confirmed by DNA sequencing (Lark Technologies).

4.2.2 Expression and Purification of C-terminal Truncated CFP-10

For protein expression, the pET28a vector containing the tCFP-10 insert was transformed into *E. coli* BL21 (DE3) (Appendix 1.1). The cells were grown in LB broth containing 40 µg/ml kanamycin until mid-log phase (optical density of 0.6 at 600 nm) and induced with IPTG to a final concentration of 0.45 mM. Cells were grown for four hours at 37 °C and 200 rpm before being harvested by centrifugation at 7,800 g for 15 minutes at 4 °C. The cell pellet was resuspended in lysis buffer (50 mM Tris, 2 mM EDTA, 0.1 % Triton X-100

(v/v), pH 8.0), to which lysozyme was added to a final concentration of 0.1 mg/ml, before incubating on a rocker at room temperature for 20 minutes. The lysed cell samples were sonicated as described for CFP-10 in section 3.2.1.2. The total cell lysate was centrifuged at 12,100 g for 15 minutes at 4 °C. The lysate supernatant containing the tCFP-10 protein was dialysed against 20 mM Tris (pH 8.0) following the addition of EDTA (1 mM) and PMSF (100 µM). The tCFP-10 sample was purified by anion exchange chromatography using a 10ml Q-sepharose column (Amersham Biosciences) and gel filtration through an Amersham HiLoad 16/60 superdex 75 gel filtration column (Section 3.2.1.2).

4.2.3 Expression and Purification of C-terminal Truncated ESAT-6

The truncated ESAT-6 (tESAT-6) protein (residues 1-84) is produced as a by product during the expression and purification of the full length ESAT-6 (Renshaw *et al.*, 2002). *E. coli* BL21 (DE3) cells containing the pET21a plasmid with full length ESAT-6 insert were grown in LB supplemented with 100 µg/ml ampicillin (Appendix 1.2). The cells were grown to mid-log phase prior to induction with IPTG (0.45 mM). Four hours post-induction the cells were harvested by centrifugation and resuspended in lysis buffer with lysozyme at a final concentration of 0.1 mg/ml. The sample was incubated on a rocker at room temperature for 20 minutes prior to sonication (Section 3.2.1.2). ESAT-6 is produced as insoluble inclusion bodies which were isolated, washed and solubilised as described in section 3.2.1.3. Proteolytic cleavage of 11 C-terminal residues from the full length ESAT-6 protein results in production of the tESAT-6 protein. Truncated ESAT-6 was separated from the full length protein by anion exchange using Q-sepharose (Amersham Biosciences), where tESAT-6 was eluted in the 100-150 mM NaCl step prior to the full length protein, which was eluted in the 150 mM NaCl wash step. Anion exchange chromatography was repeated prior to gel filtration (Section 3.2.1.3).

4.2.4 Electrospray Mass Spectrometry Sample Preparation

Electrospray mass spectrometry was performed by PNAACL (University of Leicester). The purified tCFP-10 protein sample was desalted by dialysis against dH₂O. The salt-free sample was flash frozen in liquid N₂ before freeze-drying. For mass spectrometry analysis the tCFP-10 sample was dissolved in 50 µl 0.3 % (v/v) formic acid, 50 % (w/v) acetonitrile solution to a final concentration of 10 pmol/µl.

4.2.5 Fluorescence Assays of Complex Formation between Variant CFP-10/ESAT-6 Proteins

Intrinsic fluorescence spectra of the four individual proteins (CFP-10, ESAT-6, tCFP-10 and tESAT-6) and the four possible CFP-10•ESAT-6 complexes, corresponding to CFP-10 bound to ESAT-6 (CFP-10•ESAT-6), tCFP-10 bound to ESAT-6 (tCFP-10•ESAT-6), CFP-10 bound to tESAT-6 (CFP-10•tESAT-6) and tCFP-10 bound to tESAT-6 (tCFP-10•tESAT-6) were acquired on a Perkin Elmer LS50B luminescence spectrometer. Individual protein samples were analysed at 1 µM in 25 mM Na₂HPO₄, 100 mM NaCl buffer (pH 7.5). Samples of the complexes were prepared using 1 µM CFP-10 or tCFP-10 with increasing concentrations of ESAT-6 or tESAT-6 (0 to 2.25 µM). The samples were equilibrated for 30 minutes at room temperature prior to analysis. The spectra were recorded at 20 °C with excitation at 280 nm. Fluorescence was monitored between 300 and 400 nm. The final spectra were the result of a smoothed average of 10 accumulations collected at a scan rate of 150 nm per minute, and corrected for the buffer background.

4.2.6 Binding of Full Length and Truncated CFP-10•ESAT-6 Complexes to Host Cells

4.2.6.1 Fluorescent Labelling of CFP-10•ESAT-6 Complexes

The CFP-10•ESAT-6 complexes (CFP-10•ESAT-6, tCFP-10•ESAT-6 and CFP-10•tESAT-6) were prepared by mixing equimolar solutions (25 µM) of the purified proteins at room

temperature for 30 minutes in 25 mM Na₂HPO₄, 150 mM NaCl buffer (pH 6.5). The complexes were dialysed against 25 mM Na₂HPO₄, 100 mM NaCl buffer (pH 7.5) prior to labelling with the fluorophore Alexa Fluor 546. The labelling reaction was carried out as described for the full-length CFP-10•ESAT-6 complex and the degree of labelling was calculated as per the manufacturer's instructions (Section 3.2.3).

4.2.6.2 U937 Cell Culture

Stock U937 cells were provided by Dr Jim Norman (Department of Biochemistry, University of Leicester). The U937 cells were cultured as described in section 3.2.4.2.

4.2.6.3 Fluorescence Microscopy

Monocytic U937 cells, grown in suspension, were allowed to adhere to glass coverslips pre-coated with 160 µg/ml poly-L-lysine for 20 minutes at 37 °C and 5 % CO₂. Excess cells were removed and the coverslips were washed twice with PBS. The U937 cells were exposed to 1 µM labelled complex, either CFP-10•ESAT-6, tCFP-10•ESAT-6 or CFP-10•tESAT-6 for 15 minutes at 4 °C. Any unbound complex was removed by washing with PBS. Cells were then fixed in 4 % paraformaldehyde (PFA) and permeabilised by 0.2 % (v/v) Triton X-100. Finally, coverslips were washed with PBS and dH₂O before mounting onto glass slides using ProLong Antifade Reagent (Molecular Probes). The slides were stored in the dark until dry. Fluorescence microscopy slides were observed using a Nikon TE300 inverted microscope and the images captured using an OCRA ER CCD camera (Hamamatsu) and Openlab software (Improvision), as described in Section 3.2.5.

4.2.7 Calculation of the Solvent Exposure of Tryptophan Residues

The solvent exposure of the four tryptophan residues in the full length CFP-10•ESAT-6 complex was determined from the family of 28 converged solution structures using

Molmol (Koradi *et al.*, 1996). Complexes containing C-terminal truncated variants of CFP-10 and ESAT-6 were also produced using Molmol and the probable effects on the solvent exposure of the tryptophan residues determined.

4.3 Results

4.3.1 Cloning of C-terminal Truncated CFP-10

The tCFP-10 coding sequence was amplified by PCR from the BAC Rv414 template (Brosch *et al.*, 1998). Confirmation of amplification was provided by DNA gel electrophoresis, which clearly showed a band running just below the 300 bp marker of the DNA ladder, as expected of the tCFP-10 sample (approximately 279 bp) (Figure 4.3). Following purification, the PCR product and the pET28a vector were digested with the *Nco*I and *Bam*HI enzymes for 7 hours at 37 °C to ensure complete digestion. The gel shown in figure 4.4 confirms that the tCFP-10 product was successfully ligated into the pET28a vector. The sequence integrity of the tCFP-10 insert was confirmed by DNA sequencing of the ligated plasmid (Lark Technologies) (Appendix 4.13).

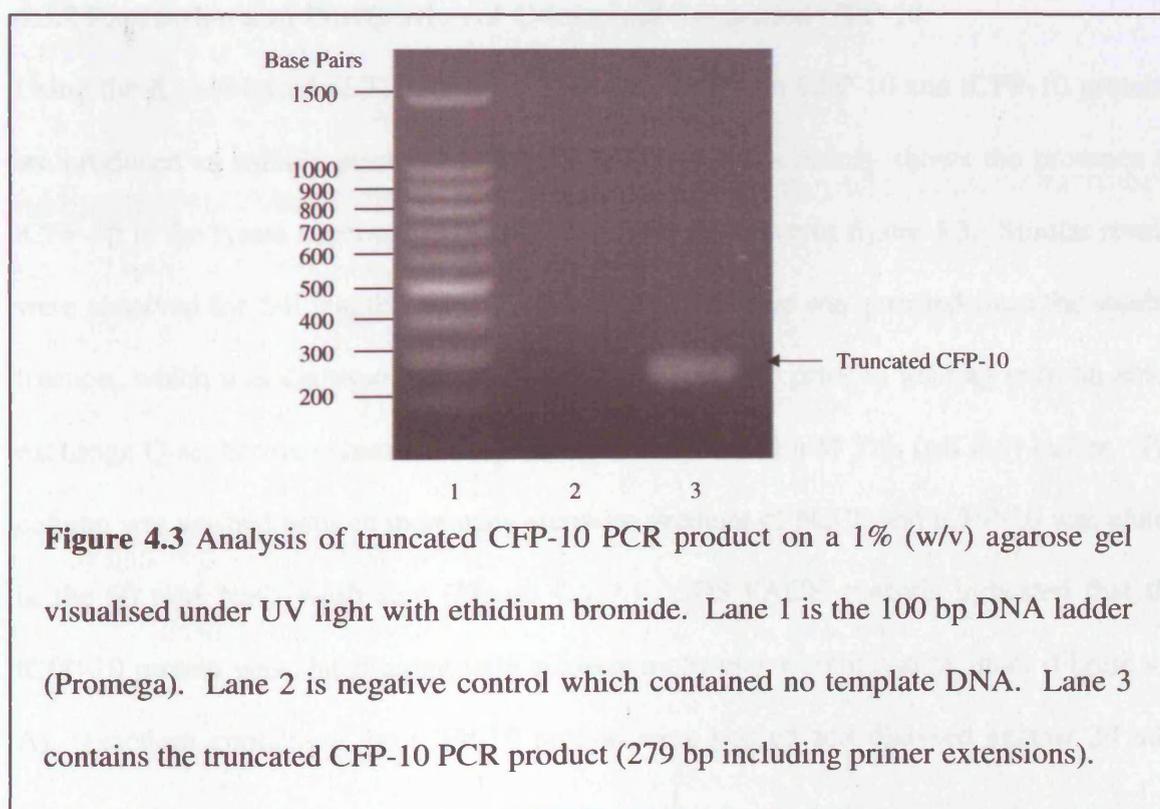


Figure 4.3 Analysis of truncated CFP-10 PCR product on a 1% (w/v) agarose gel visualised under UV light with ethidium bromide. Lane 1 is the 100 bp DNA ladder (Promega). Lane 2 is negative control which contained no template DNA. Lane 3 contains the truncated CFP-10 PCR product (279 bp including primer extensions).

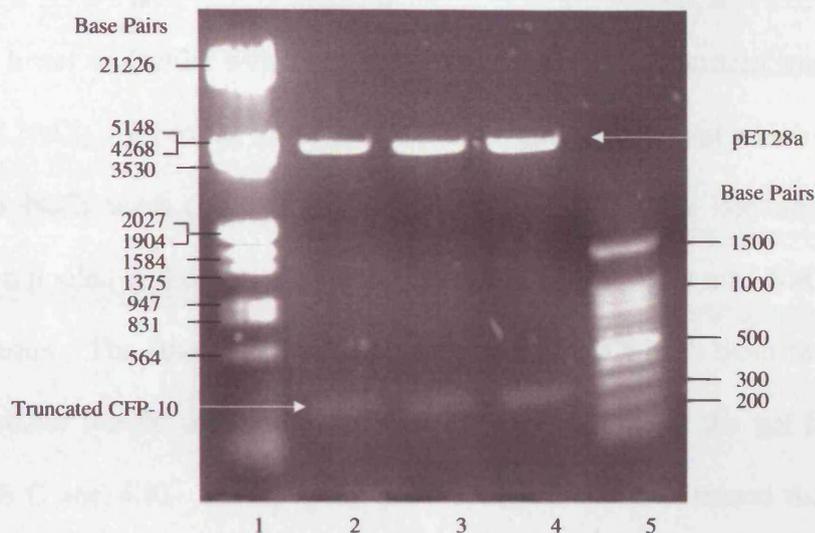


Figure 4.4 A 1% (w/v) agarose gel showing dual restriction digests of the pET28a vector containing the truncated CFP-10 PCR insert. The pET28a plasmid was digested using *NcoI* and *BamHI* restriction enzymes (Promega). Lane 1 is the λ DNA/*EcoRI* + *HindIII* marker and lane 5 is the 100 bp marker (Promega). Lanes 2, 3 and 4 contain the pET28a digested samples with truncated CFP-10 inserts.

4.3.2 Expression and Purification of C-terminal Truncated CFP-10

Using the *E. coli* based pET28a system both the full length CFP-10 and tCFP-10 proteins are produced as soluble products and SDS-PAGE analysis clearly shows the presence of tCFP-10 in the lysate supernatant (soluble fraction) as shown in figure 4.5. Similar results were observed for full length CFP-10. The tCFP-10 protein was purified from the soluble fraction, which was dialysed against 20 mM Tris (pH 8.0) prior to loading onto an anion exchange Q-sepharose column pre-equilibrated with the 20 mM Tris (pH 8.0) buffer. The column was washed with an increasing stepwise gradient of NaCl, and tCFP-10 was eluted in the 50 mM NaCl wash step (Figure 4.6 A). SDS-PAGE analysis indicated that the tCFP-10 protein was eluted along with a lower molecular weight contaminant (Figure 4.7 A). Fractions containing the tCFP-10 protein were pooled and dialysed against 20 mM

Piperazine (pH 5.8) prior to re-loading onto the equilibrated Q-sepharose column to remove the lower molecular weight contaminant. The tCFP-10 protein was eluted between 0 to 40 mM NaCl, prior to the lower molecular weight contaminant which is eluted later in the 40 mM NaCl wash (Figures 4.6 B and 4.7 B). Fractions containing the tCFP-10 protein were pooled and dialysed against 25 mM Na₂HPO₄, 150 mM NaCl (pH 6.5) prior to gel filtration. The final polishing step separated the tCFP-10 from trace levels of the lower molecular weight contaminant which ran slower through the gel filtration column (Figures 4.6 C and 4.8). Electrospray mass spectroscopy determined that the molecular weight of the tCFP-10 protein was 9082.97 Da (Figure 4.9). This agrees with the predicted mass of 9082.9 Da which corresponds to the tCFP-10 protein minus the N-terminal methionine.

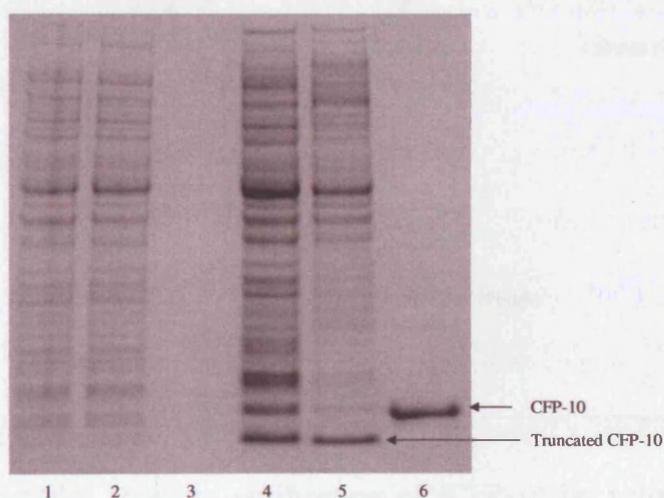


Figure 4.5 A Coomassie stained SDS-PAGE gel showing the expression of the tCFP-10 protein using the pET28a system. Lane 1 is the pre-induction whole cell sample and lane 2 is the four hour post-induction whole cell sample. Lane 3 is the culture supernatant. Lane 4 is the total cell lysate and lane 5 is the lysate supernatant. Lane 6 is a full length CFP-10 control. The gel clearly shows the presence of tCFP-10 in the lysate supernatant (soluble fraction) in lane 5.

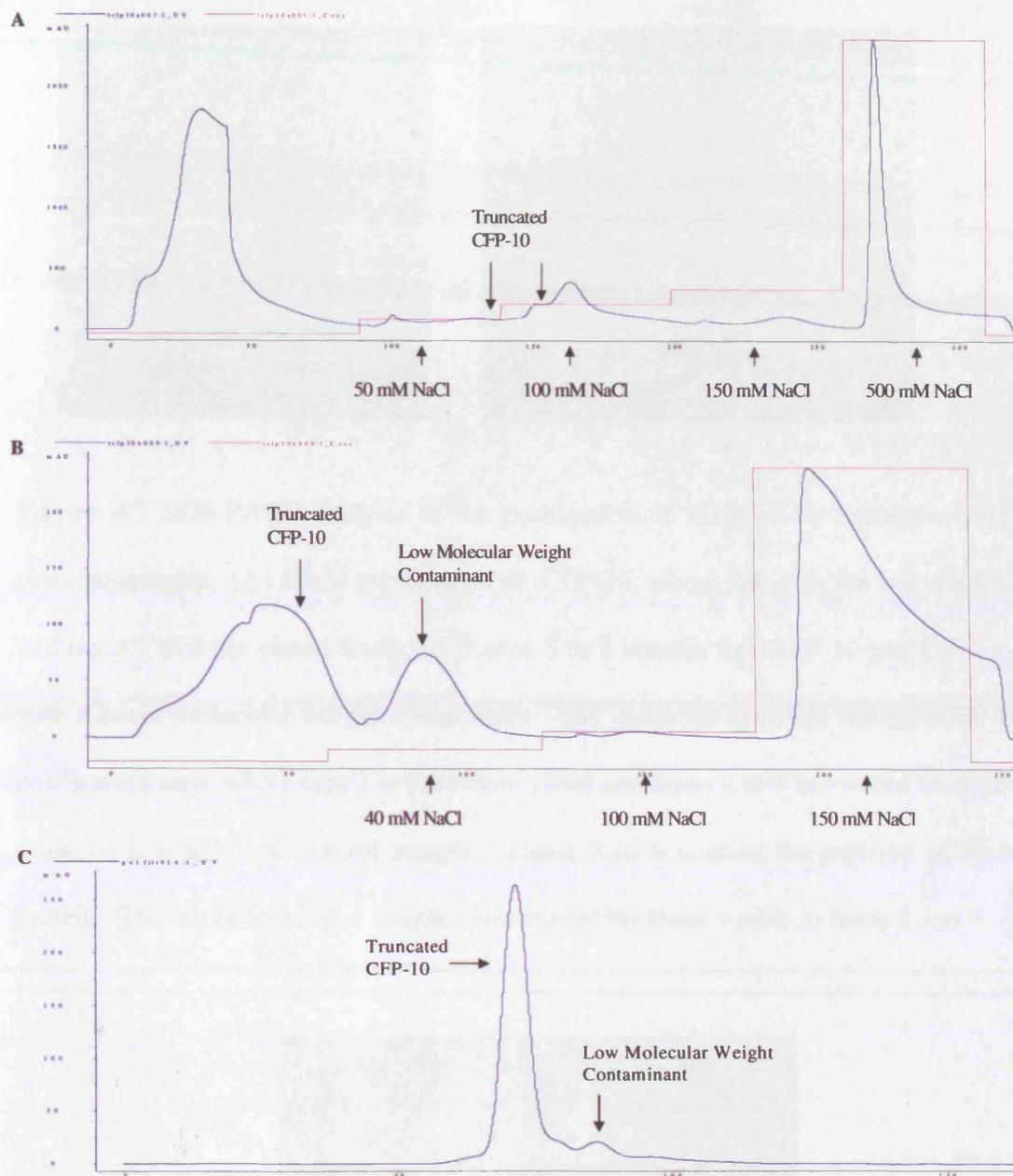


Figure 4.6 Elution profiles showing purification of tCFP-10 by anion exchange chromatography. (A) In the first step the Q-sepharose column is equilibrated with 20 mM Tris (pH 8.0). The tCFP-10 protein is eluted in the 50-100 mM NaCl wash. (B) The second anion exchange purification step is used to separate the tCFP-10 protein from a lower molecular weight contaminant. Truncated CFP-10 is eluted prior to the contaminant, which is eluted in the 40 mM NaCl wash. The base buffer used in this step was 20 mM Piperazine (pH 5.8). (C) Gel filtration purification of the tCFP-10 protein removes any remaining lower molecular weight contamination.

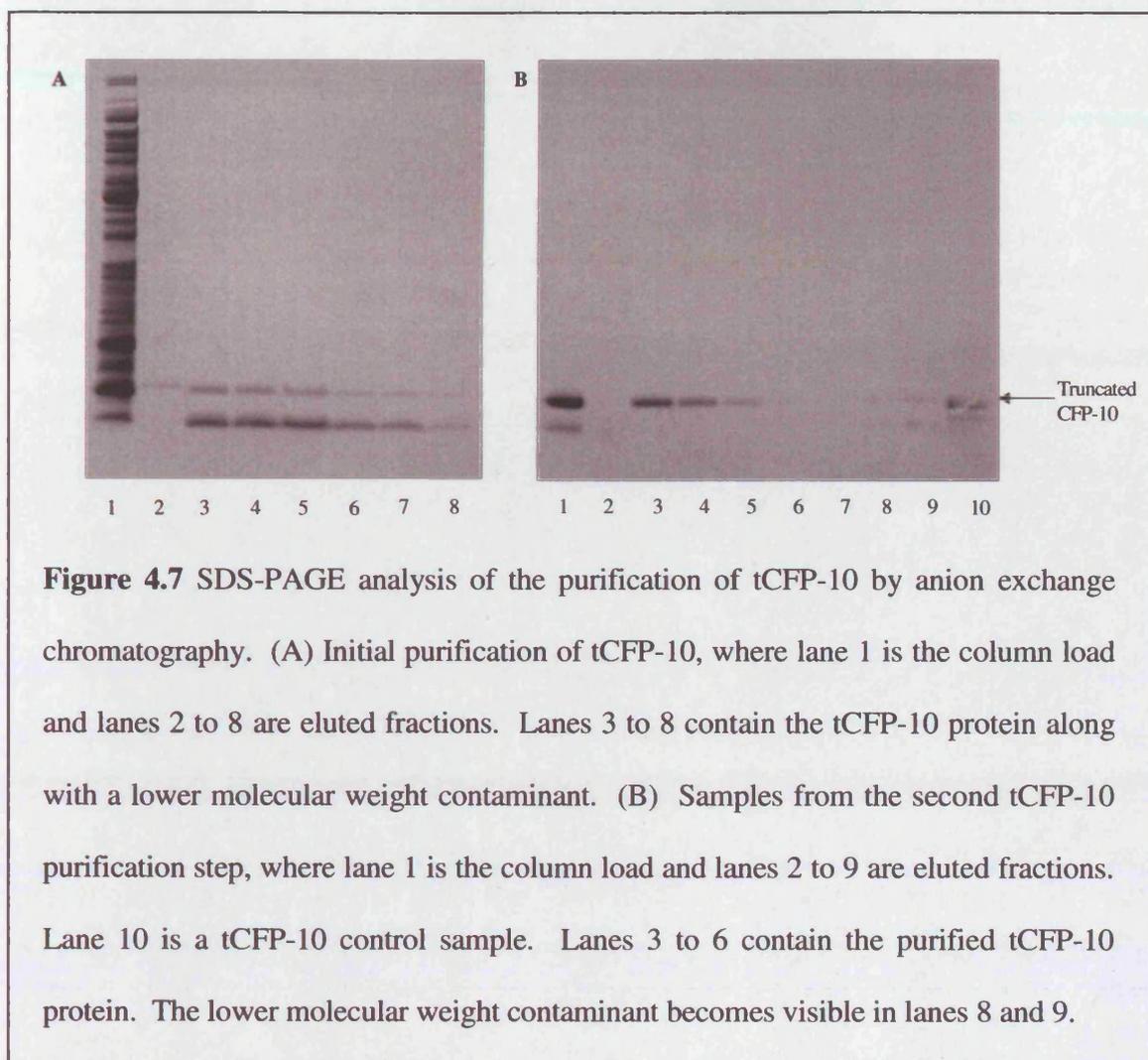


Figure 4.7 SDS-PAGE analysis of the purification of tCFP-10 by anion exchange chromatography. (A) Initial purification of tCFP-10, where lane 1 is the column load and lanes 2 to 8 are eluted fractions. Lanes 3 to 8 contain the tCFP-10 protein along with a lower molecular weight contaminant. (B) Samples from the second tCFP-10 purification step, where lane 1 is the column load and lanes 2 to 9 are eluted fractions. Lane 10 is a tCFP-10 control sample. Lanes 3 to 6 contain the purified tCFP-10 protein. The lower molecular weight contaminant becomes visible in lanes 8 and 9.

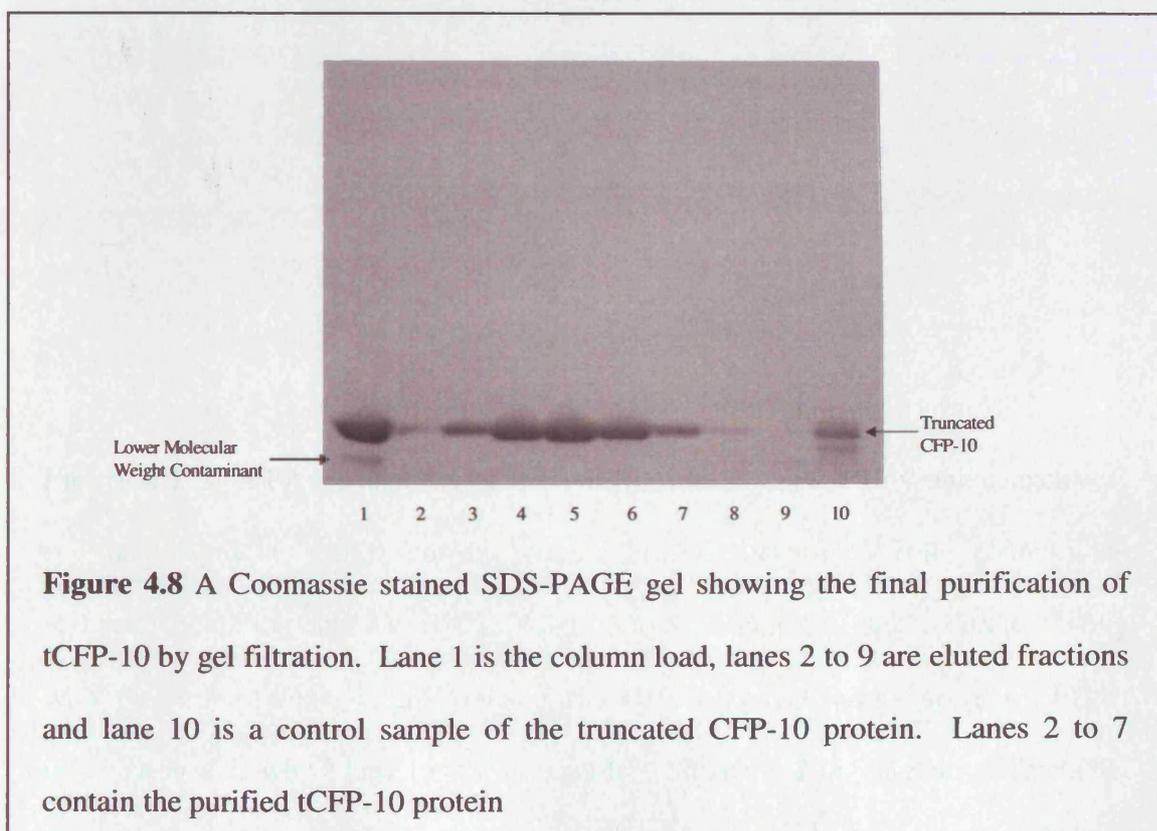


Figure 4.8 A Coomassie stained SDS-PAGE gel showing the final purification of tCFP-10 by gel filtration. Lane 1 is the column load, lanes 2 to 9 are eluted fractions and lane 10 is a control sample of the truncated CFP-10 protein. Lanes 2 to 7 contain the purified tCFP-10 protein

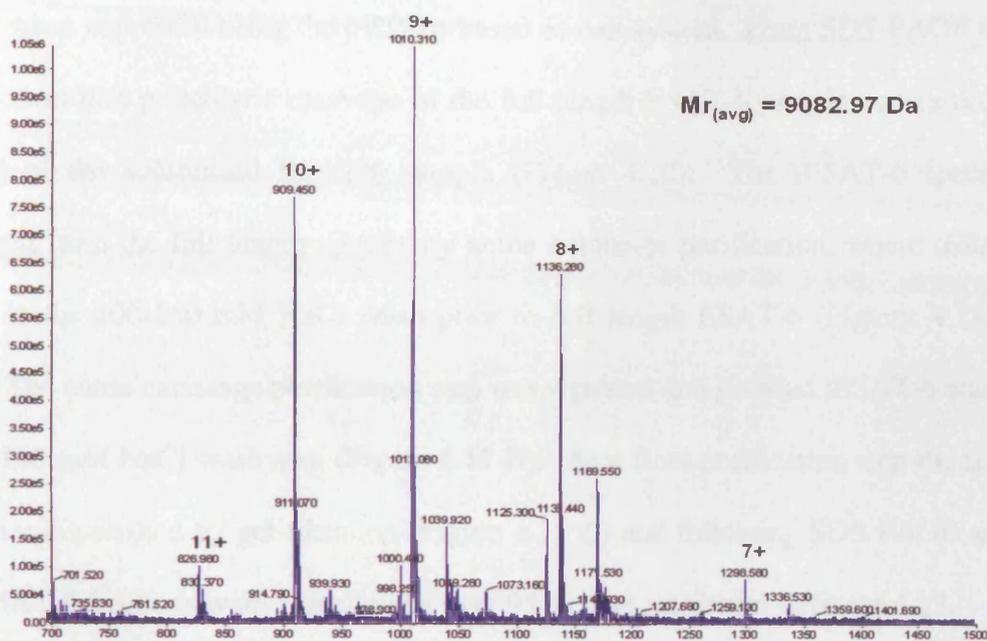
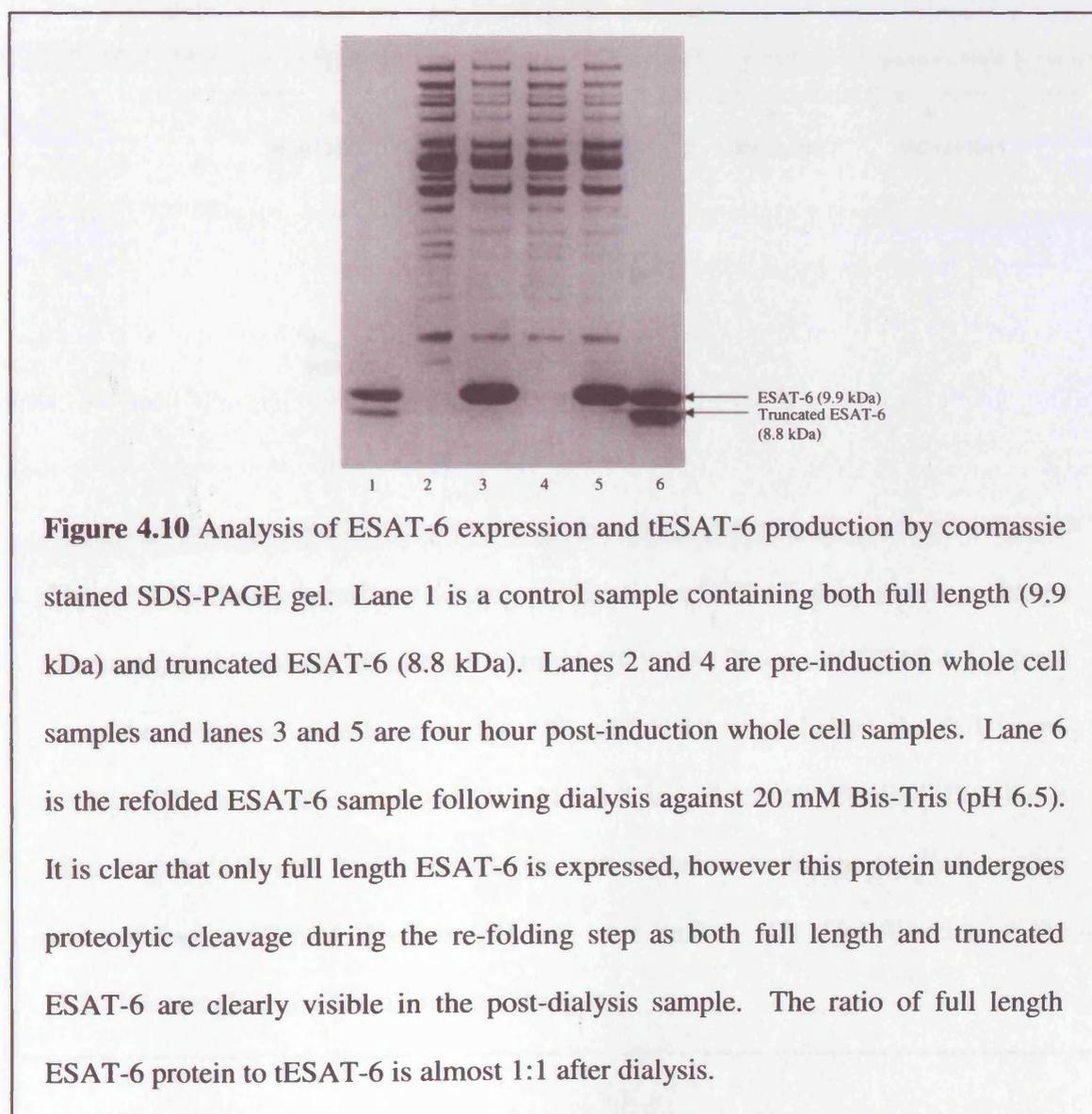


Figure 4.9 Electrospray mass spectroscopy mass:charge envelope for C-terminal truncated CFP-10. The data indicate that the molecular weight of the purified tCFP-10 protein was 9082.97 Da, which corresponds to the truncated CFP-10 protein lacking the N-terminal methionine.

4.3.3 Expression and Purification of C-terminal Truncated ESAT-6

A truncated variant of ESAT-6, lacking 11 residues from the C-terminus is naturally produced by *M. tuberculosis* H37Rv (Okkels *et al.*, 2004). Likewise, expression and purification of the full length protein using the *E. coli* pET21a-based system often results in the production of a truncated ESAT-6 species (Renshaw *et al.*, 2002). This process is particularly evident if the protease inhibitors EDTA and PMSF are omitted from the refolding buffer, indicating that the tESAT-6 protein is the result of proteolytic cleavage of the full length ESAT-6 protein. Two major species of ESAT-6 were previously identified by electrospray mass spectroscopy, the full length protein (9.9 kDa) and the truncated species, lacking 11 residues from the C-terminal (8.8 kDa) (Renshaw *et al.*, 2002).

As described in section 3.3.1.2, full length ESAT-6 is produced as insoluble inclusion bodies when expressed using the pET21a-based *E. coli* system. From SDS-PAGE analysis it is evident that proteolytic cleavage of the full length ESAT-6 protein occurs during re-folding of the solubilised ESAT-6 sample (Figure 4.10). The tESAT-6 species was separated from the full length species by anion exchange purification, where tESAT-6 is eluted in the 100-150 mM NaCl wash prior to full length ESAT-6 (Figures 4.11 A and 4.12). The anion exchange purification step was repeated and purified tESAT-6 was eluted in the 150 mM NaCl wash step (Figure 4.11 B). As a final purification step the tESAT-6 sample was polished by gel filtration (Figure 4.11 C) and following SDS-PAGE analysis, the tESAT-6 fractions were judged to be over 95 % pure, as shown in figure 4.13.



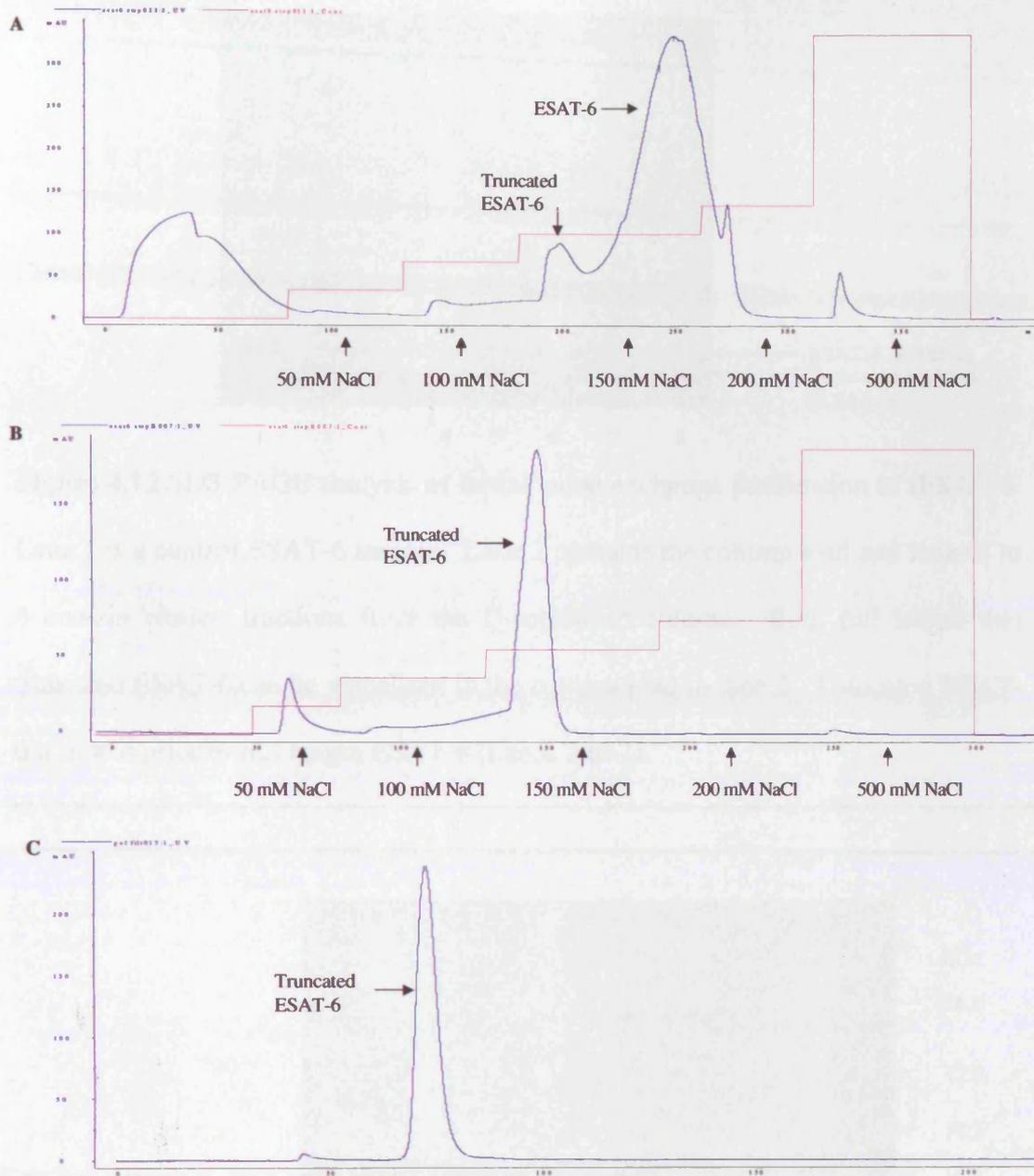


Figure 4.11 Elution profiles showing purification of tESAT-6 by anion exchange chromatography (A and B) and gel filtration (C). (A) Truncated ESAT-6 is eluted from the Q-sepharose column in the 150 mM NaCl wash before the full length protein. The base buffer used was 20 mM Bis-Tris (pH 6.5) (B) Purified tESAT-6 is eluted in the 150 mM NaCl wash in the second anion exchange purification step using the same 20 mM Bis-Tris (pH 6.5) base buffer. (C) Gel filtration of the tESAT-6 sample is used as a final polishing step.

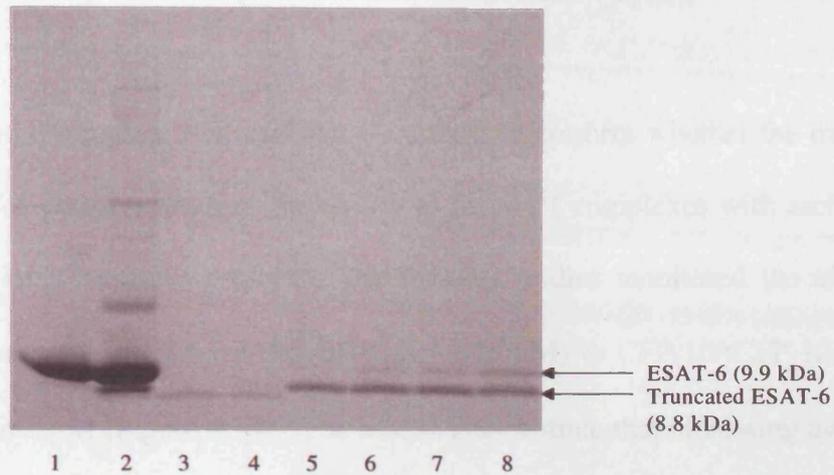


Figure 4.12 SDS-PAGE analysis of initial anion exchange purification of tESAT-6.

Lane 1 is a control ESAT-6 sample. Lane 2 contains the column load and lanes 3 to 8 contain elution fractions from the Q-sepharose column. Both full length and truncated ESAT-6 can be visualised in the column load in lane 2. Truncated ESAT-6 is eluted prior to full length ESAT-6 (Lanes 5 to 7).

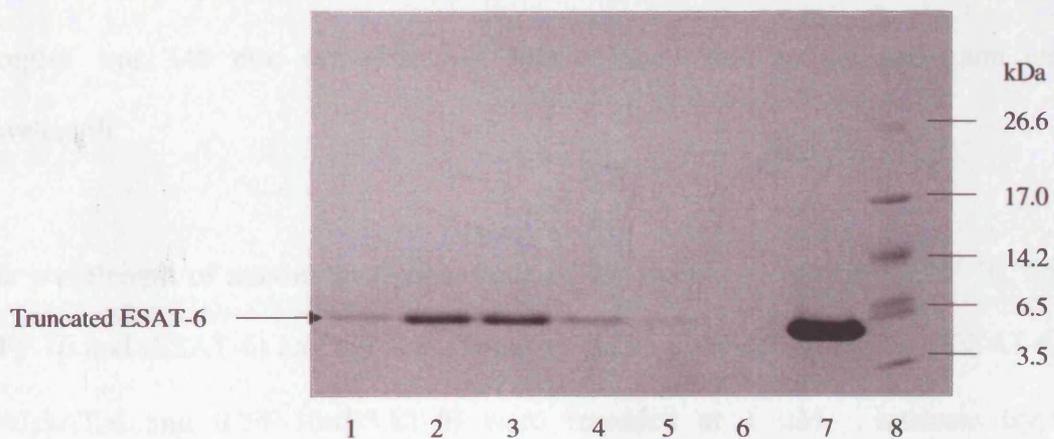


Figure 4.13 Analysis of purified tESAT-6 by a coomassie stained SDS-PAGE gel.

Fractions from the gel filtration column were loaded in lanes 1 to 6. Lane 7 is a tESAT-6 control sample. Lane 8 is the ultra low molecular weight marker (Sigma). The fractions in lanes 1 to 4 contain purified tESAT-6 protein.

4.3.4 Fluorescence Assays of Complex Formation between Variant CFP-10/ESAT-6 Proteins

Changes in intrinsic tryptophan fluorescence were used to confirm whether the truncated CFP-10 and ESAT-6 proteins retained the ability to form 1:1 complexes with each other, or with their full length partner protein. The binding studies monitored the effect of increasing the molar ratio of ESAT-6/tESAT-6 (0 - 2.25 μ M) to CFP-10/tCFP-10, which was kept constant at 1 μ M (Figure 4.14). The results demonstrate that increasing the molar concentration of ESAT-6 has a similar effect with CFP-10 as with tCFP-10. In both cases the shortest wavelength (344 nm) is recorded at a molar ratio of 1:1 ESAT-6 to CFP-10 or tCFP-10, indicating that full length ESAT-6 is capable of forming a 1:1 complex with both CFP-10 and tCFP-10. The studies also show that tESAT-6 and CFP-10 are able to form a 1:1 complex. In contrast, the binding curve for the dual truncated complex is significantly different to those obtained for the CFP-10•tESAT-6, tCFP-10•ESAT-6 and CFP-10•tESAT-6 complexes. The shortest wavelength observed for the tCFP-10•tESAT-6 complex was 348 nm, demonstrating only a small shift in the maximum emission wavelength.

The wavelength of maximum fluorescence of the individual proteins (CFP-10, ESAT-6, tCFP-10 and tESAT-6) and the 1:1 complexes (CFP-10•ESAT-6, tCFP-10•ESAT-6, CFP-10•tESAT-6 and tCFP-10•tESAT-6) were recorded at 1 μ M. Intrinsic tryptophan fluorescence spectra recorded for CFP-10, ESAT-6, tCFP-10 and tESAT-6 demonstrate that the wavelength of maximum emission for the individual proteins range between 352 nm to 354 nm (Figure 4.15), consistent with the tryptophan residues being in a solvent exposed position. The fluorescence maximum observed for the full length 1:1 CFP-10•ESAT-6 complex was blue shifted to 344 nm, indicating that at least one of the four tryptophan residues has become significantly less exposed to the solvent (Figure 4.15 A).

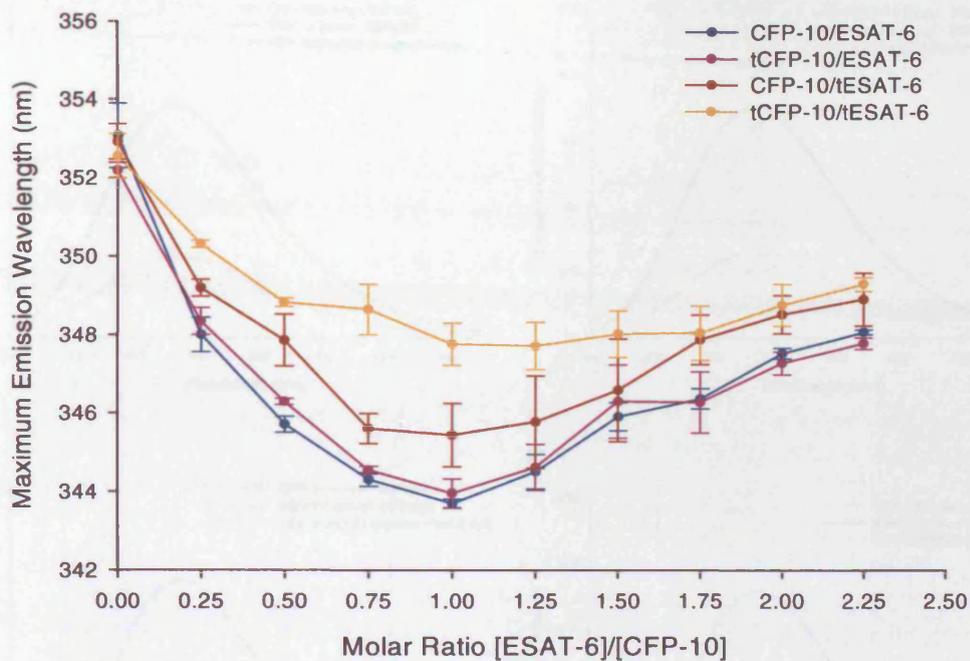


Figure 4.14 Fluorescence-based assay of complex formation by full length and truncated CFP-10 and ESAT-6 proteins. The graph shows the changes observed in the wavelength of maximum intrinsic fluorescence on increasing the molar ratio of ESAT-6 or tESAT-6 (0-2.25 μ M) to CFP-10 or tCFP-10 (1 μ M). The data indicates that the following 1:1 complexes can be formed CFP-10•ESAT-6, tCFP-10•ESAT-6, CFP-10•tESAT-6. The data obtained for binding of the two truncated species is less convincing, with the shift in fluorescence being significantly reduced (from 353 nm to 348 nm), suggesting that complex formation between the two truncated proteins is significantly perturbed. The data shown are based on results from at least two independent assays and error bars represent the standard deviation.

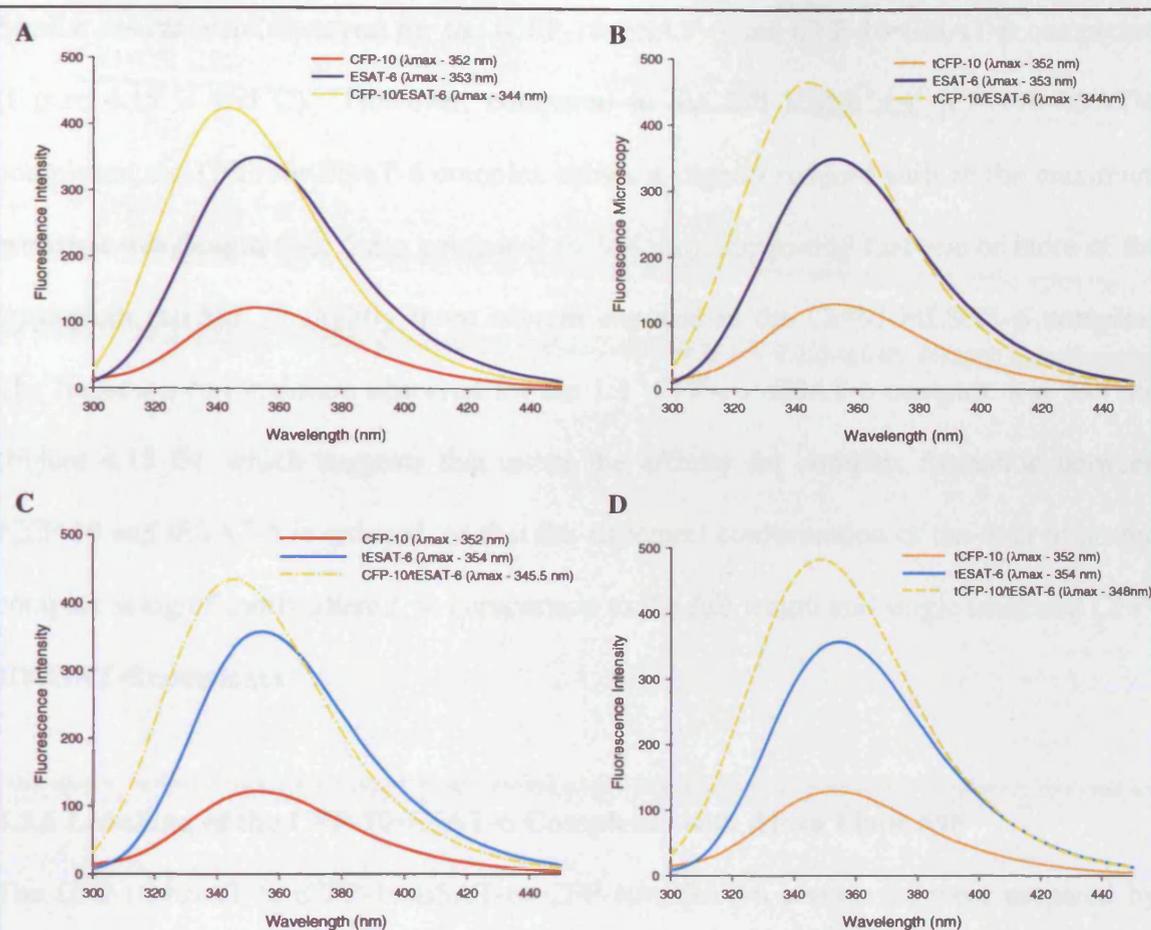


Figure 4.15 Fluorescence emission spectra of 1 μ M individual CFP-10, ESAT-6, tCFP-10 and tESAT-6 proteins and their complexes (CFP-10•ESAT-6, tCFP-10•ESAT-6, CFP-10•tESAT-6 and tCFP-10•tESAT-6). (A) The wavelength of maximum fluorescence emission for CFP-10 and ESAT-6 is 352 and 353 nm, respectively, suggesting that the tryptophan residues are fully solvent exposed. The 1:1 CFP-10•ESAT-6 complex is blue shifted to 344 nm, indicating that one or more of the tryptophan residues has become partially or fully buried upon complex formation. Similar results are observed for single truncated complexes tCFP-10•ESAT-6 (B) and CFP-10•tESAT-6 (C). The maximum fluorescence emission observed for the dual truncated complex (tCFP-10•tESAT-6) (D) is 348 nm, suggesting that the complex formed between tCFP-10 and tESAT-6 is structurally different to those observed for the full length and singularly truncated complexes.

Similar results were observed for the tCFP-10•ESAT-6 and CFP-10•tESAT-6 complexes (Figure 4.15 B and C). However, compared to the full length and tCFP-10•ESAT-6 complexes, the CFP-10•tESAT-6 complex shows a slightly reduced shift in the maximum emission wavelength (345.5 nm compared to 344 nm), suggesting that one or more of the tryptophan residues is slightly more solvent exposed in the CFP-10•tESAT-6 complex. The fluorescence maximum observed for the 1:1 tCFP-10•tESAT-6 complex was 348 nm (Figure 4.15 D), which suggests that either the affinity for complex formation between tCFP-10 and tESAT-6 is reduced, or that the structural conformation of the dual truncated complex is significantly altered, in comparison to the full length and single truncated CFP-10/ESAT-6 complexes

4.3.5 Labelling of the CFP-10•ESAT-6 Complexes with Alexa Fluor 546

The CFP-10•ESAT-6, tCFP-10•ESAT-6, CFP-10•tESAT-6 complexes were prepared by mixing equimolar concentrations (25 μ M) of CFP-10 or tCFP-10 with either ESAT-6 or tESAT-6 for 30 minutes at room temperature. The complexes were labelled with Alexa Fluor 546 at pH 7.5, with the aim of selectively labelling the N-termini of both proteins within the complex. The degree of labelling for each of the complexes was calculated to be: 1.00 for the CFP-10•tESAT-6 complex, 1.38 for the tCFP-10•ESAT-6 complex (Figure 4.16) and 1.57 for the full length complex (Figure 3.13).

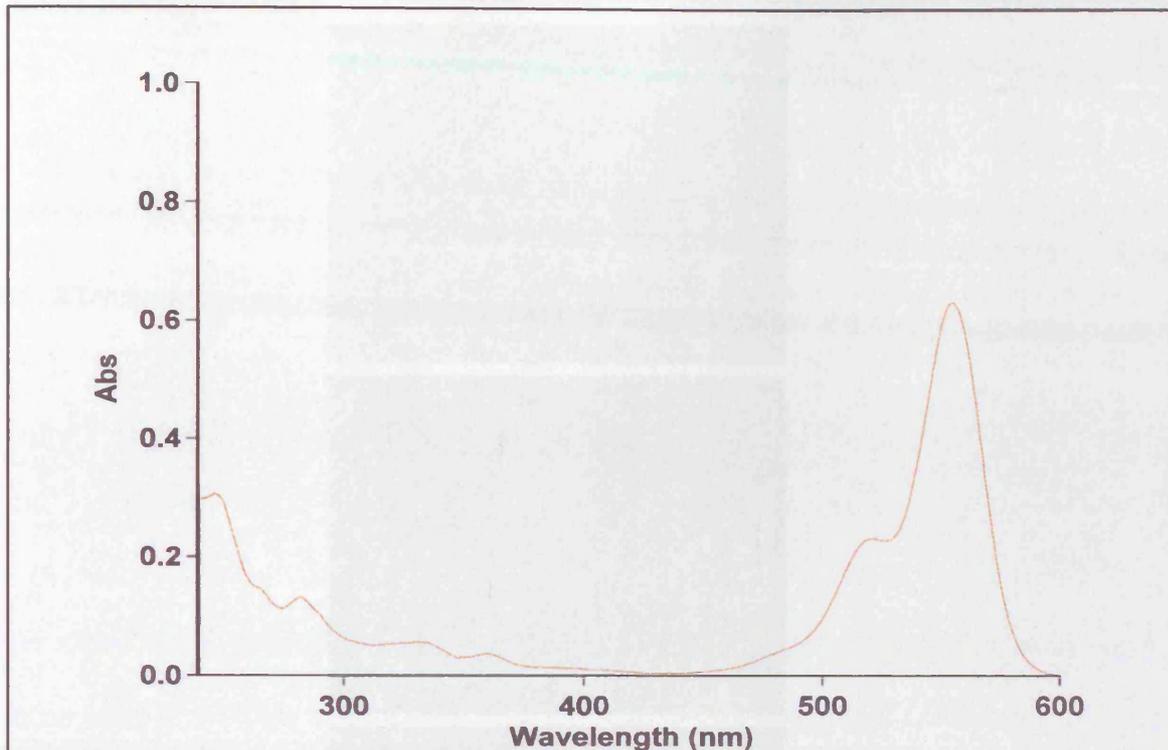


Figure 4.16 Absorbance spectra of the Alexa Fluor 546 labelled tCFP-10•ESAT-6 complex. For the example shown the degree of labelling was calculated to be 1.38.

4.3.6 Fluorescence Microscopy Assays of Variant CFP-10•ESAT-6 Complexes Binding to U937 Cells

The fluorescence labelling of cells exposed to the CFP-10•ESAT-6 complex is very similar to that observed for the full length CFP-10•ESAT-6 complex demonstrating that both of these complexes bind to the surface of the U937 cells (Figure 4.17 A-B). However, the labelling observed for the tCFP-10•ESAT-6 complex is dramatically reduced in comparison to the full length complex (Figure 4.17). This suggests that removal of the C-termini of CFP-10 reduces binding of the labelled tCFP-10•ESAT-6 to the monocyte cell surface, providing clear evidence that this region forms part of the binding site for the cell surface receptor.

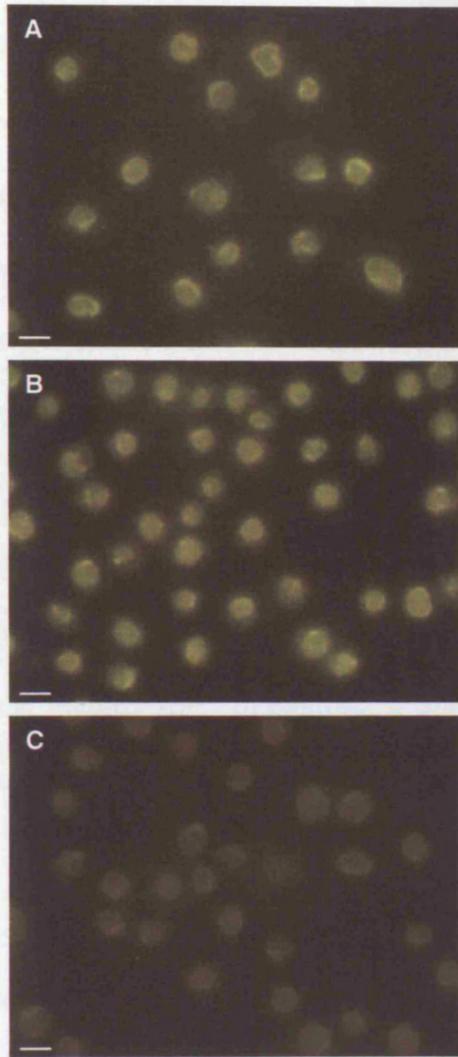


Figure 4.17 Reduced binding of the tCFP-10•ESAT-6 complex to the surface of U937 monocyte cells. Panel (A) shows the binding of the full length CFP-10•ESAT-6 complex to the surface of U937 cells (250 ms exposure). Panels (B) and (C) are the result of incubating U937 cells with the CFP-10•tESAT-6 complex and the tCFP-10•ESAT-6 complex respectively (200 ms exposure). The fluorescence intensity observed for the CFP-10•ESAT-6 complex is similar to that observed for the full length complex. In contrast, the fluorescence intensity of U937 cells exposed to the tCFP-10•ESAT-6 complex is significantly reduced in comparison to the CFP-10•ESAT-6 and CFP-10•tESAT-6 complexes, indicating that the C-terminal of CFP-10 is likely to be involved in interactions with the surface of host cells. Size bars shown are 10 μ M.

4.4 Discussion

The fluorescence microscopy studies, described in the previous chapter, demonstrated that the CFP-10•ESAT-6 complex specifically binds to the surface of monocyte cells. The region of the complex involved in this interaction is unknown, however, the striking nature of the flexible C-terminal arms of the complex suggested a possible role in binding. To investigate this possibility, truncated CFP-10 and ESAT-6 proteins were complexed with their full length partners to form various CFP-10•ESAT-6 complexes for use in fluorescence microscopy assays. The tCFP-10 construct was designed to remove 14 residues from the C-terminus. This design, based on the structure of the CFP-10•ESAT-6 complex, aimed to retain binding of tCFP-10 to the ESAT-6 or tESAT-6 partner protein but eliminate the potential binding site. In contrast, a truncated ESAT-6 protein, lacking 11 residues from the C-terminus, is produced as a result of proteolytic cleavage of the full length protein.

Prior to fluorescence microscopy assays, complex formation between the full length and truncated CFP-10/ESAT-6 proteins was characterised by fluorescence spectroscopy. In the fluorescence-based assays described here, the changes in intrinsic tryptophan emission observed for complex formation between the full length CFP-10 and ESAT-6 are consistent with previous studies by Renshaw *et al.* (2002), who showed that CFP-10 and ESAT-6 form a 1:1 complex. The fluorescence-derived binding curves, shown in figure 4.14, clearly demonstrate that the shortest wavelength of maximum fluorescence emission was attained at a 1:1 molar ratio of ESAT-6 to CFP-10. Similar results were observed for the tCFP-10•ESAT-6 and CFP-10•tESAT-6 complexes (Figure 4.14), clearly indicating that both tCFP-10 and tESAT-6 are able to form 1:1 complexes with their full length partner proteins.

The binding curves for interactions between CFP-10/ESAT-6 and tCFP-10/ESAT-6 were highly similar and demonstrate that both full length CFP-10 and the C-terminal truncated protein form equivalent 1:1 complexes with ESAT-6 (Figure 4.14). This suggests that removal of the 14 residues from the C-terminus of CFP-10 has no obvious effect on complex formation with ESAT-6, as expected from analysis of the structure of the CFP-10•ESAT-6 complex. Likewise, removal of the C-terminal 11 residues from ESAT-6 does not significantly affect complex formation with full length CFP-10. However, based on the binding curve observed for the dual truncated complex, removal of both C-terminal regions has a considerable effect on the structure of the complex and/or the affinity for complex formation between the two truncated proteins (Figure 4.14).

The wavelength of maximum fluorescence emission reported for the individual proteins was: 352nm for both CFP-10 and tCFP-10, 353 nm for ESAT-6 and 354 nm for tESAT-6 (Figure 4.15). The results indicate that the single tryptophan of CFP-10/tCFP-10 (Trp₄₃) and the three tryptophan residues of ESAT-6/tESAT-6 (Trp₆, Trp₄₃, Trp₅₈) are fully solvent exposed (Figure 4.15). The fluorescence maximum for both the full length CFP-10•ESAT-6 complex and the tCFP-10•ESAT-6 complex was reported to be 344 nm (Figure 4.15 A and B), suggesting that at least one of the four tryptophan residues has moved to a less polar environment. The results for the CFP-10•tESAT-6 complex indicates a fluorescence emission maximum of 345.5 nm which is slightly red shifted in comparison to the CFP-10•ESAT-6 and tCFP-10•ESAT-6 complexes (Figure 3.15 A-C). However, a more significant change is observed for the dual truncated protein complex. The fluorescence maximum reported for the tCFP-10•tESAT-6 complex was 348 nm, suggesting that the tryptophan residues of this complex are more solvent exposed than the full length and single truncated CFP-10/ESAT-6 complexes.

The structure of the CFP-10•ESAT-6 complex, as shown in figure 4.18, clearly shows that Trp₅₈ of ESAT-6 is fully buried within the hydrophobic core of the CFP-10•ESAT-6 protein complex. Both Trp₄₃ of ESAT-6 and Trp₄₃ of CFP-10 are partially buried within the complex, and Trp₆ positioned at the N-terminus of ESAT-6 is the most solvent exposed of the four tryptophan residues (Figure 4.18 and Table 4.1). Analysis based on the structure of the CFP-10•ESAT-6 complex suggests that removal of 14 residues from the CFP-10 C-terminus (orange in figure 4.18) or loss of the 11 C-terminal ESAT-6 residues (cyan) should have no effect on the solvent exposure of any of the four tryptophan residues (yellow) within the complex (Table 4.1). This strongly suggests that the changes observed in intrinsic tryptophan fluorescence for the complex formed from tCFP-10 and tESAT-6 is the result of a conformational change in the complex and/or a weaker binding affinity, which precluded it from the further studies of binding to the surface of host cells.

The U937 cells incubated with the labelled CFP-10•tESAT-6 complex show a similar level of fluorescence to that observed for the full length complex (Figure 4.17 A and B), indicating that removal of the ESAT-6 C-terminal has no effect on binding to the surface of monocyte cells. However, the fluorescence labelling of cells exposed to the tCFP-10•ESAT-6 complex is significantly reduced in comparison (Figure 4.17 C), suggesting that the C-terminal arm of CFP-10 forms part of the host cell receptor binding site. The significant reduction of fluorescence labelling observed with cells exposed to the tCFP-10•ESAT-6 complex also confirms the presence of a specific receptor for the full length complex, supporting the microscopy studies described in the previous chapter.

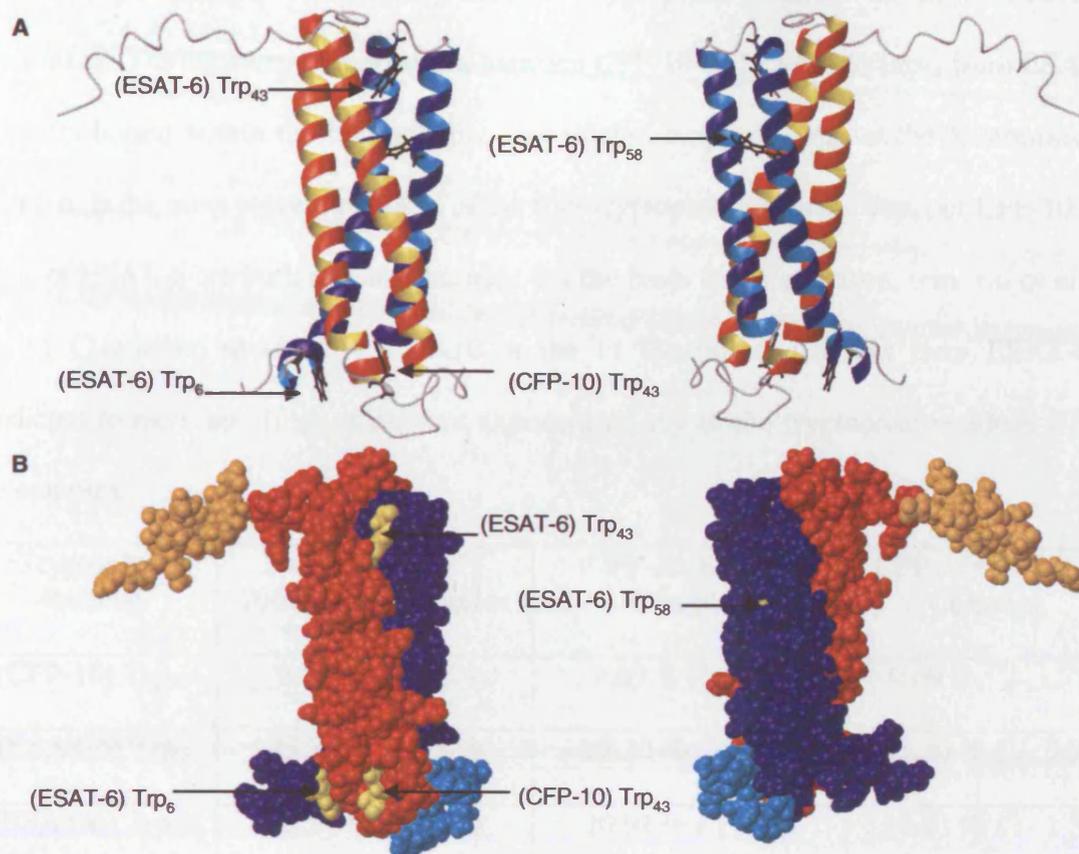


Figure 4.18 Solvent accessibility of tryptophan residues in the CFP-10•ESAT-6 complex. CFP-10 is shown in red and ESAT-6 in blue. (A) Ribbon representation of the four helix bundle structure adopted by the CFP-10•ESAT-6 complex. The complex contains four tryptophan residues, one in CFP-10 (Trp₄₃) and three in ESAT-6 (Trp₆, Trp₄₃ and Trp₅₈), which are highlighted in black. (B) A space filled view of the complex, with the 14 C-terminal residues of CFP-10 removed to produce tCFP-10 highlighted in orange and the 11 C-terminal residues removed in tESAT-6 shown in cyan. The tryptophan residues are highlighted in yellow. From the structures shown it is clear that one of the tryptophan residues is fully buried in the hydrophobic core of the protein (Trp₅₈). The three remaining residues are partially buried, with Trp₆ being the most solvent exposed.

Table 4.1 Calculated solvent exposure of tryptophan residues in CFP-10/ESAT-6 complexes. During complex formation between CFP-10 and ESAT-6, Trp₅₈ from ESAT-6 is totally buried within the hydrophobic core of the complex. Trp₆, at the N-terminus of ESAT-6, is the most solvent exposed of the four tryptophan residues. Trp₄₃ of CFP-10 and Trp₄₃ of ESAT-6 are both partially buried. On the basis of the structure, removal of either the 14 C-terminal residues of CFP-10 or the 11 C-terminal residues from ESAT-6 is predicted to have no effect on solvent exposure of any of the tryptophan residues within the complex.

Tryptophan Residue	Full-length CFP-10•ESAT-6 Complex	tCFP-10•ESAT-6 Complex	CFP-10•tESAT-6 Complex
(CFP-10) Trp ₄₃	9.69 % (\pm 1.7)	9.69 % (\pm 1.7)	9.69 % (\pm 1.7)
(ESAT-6) Trp ₆	23.30 % (\pm 5.9)	23.30 % (\pm 5.9)	23.30 % (\pm 5.9)
(ESAT-6) Trp ₄₃	10.91 % (\pm 1.3)	10.91 % (\pm 1.3)	10.91 % (\pm 1.3)
(ESAT-6) Trp ₅₈	0.04 % (\pm 0.05)	0.04 % (\pm 0.05)	0.04 % (\pm 0.05)

In conclusion, the fluorescence-based binding assays reported here clearly demonstrate that C-terminal truncated CFP-10 and ESAT-6 proteins retain the ability to form 1:1 complexes with their full length partner proteins, similar to that observed for CFP-10 and ESAT-6. However, the dual truncated complex was eliminated from further study as analysis of this complex suggested either weak binding affinity between the two truncated proteins, or that the structure of the complex is significantly perturbed. The fluorescence microscopy studies described in this chapter provide clear evidence that the C-terminal arm of CFP-10 is necessary to interact with the host cell receptor present on the surface of monocyte cells, and although the same observation is not made for the C-terminal of the ESAT-6 protein, this conserved region may still play a role in function of the CFP-10•ESAT-6 complex. Also, the reduced binding of the tCFP-10•ESAT-6 complex to the surface of monocyte

cells confirms the presence of a specific target receptor for the full length CFP-10•ESAT-6 complex on the host cell surface.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

5.1.1 Interactions between Members of the CFP-10/ESAT-6 Family

The reliable detection of complex formation between CFP-10 and ESAT-6 using the yeast two-hybrid system (Fields and Song, 1989, Chien *et al.*, 1991) introduced a convenient method to study complex formation between other members of the CFP-10/ESAT-6 protein family. The results, detailed in section 2.3.2, clearly demonstrate that the genome pairs Rv0287/Rv0288 and Rv3019c/Rv3020c form analogous heterodimeric complexes to those described for the CFP-10/ESAT-6 pair. Recent fluorescence-based binding assays carried out in the group have confirmed preliminary data obtained by collaborators at the University of Kent, and conclusively shows that Rv0287/Rv0288 form a tight 1:1 complex (Lightbody *et al.*, 2004). Further yeast two-hybrid studies provided no evidence that either CFP-10 or ESAT-6 interact with any other members of the CFP-10/ESAT-6 members, however, complex formation between the non-genome partners Rv0287/Rv3019c and Rv0288/Rv3020c was detected. The results of western-western blotting and print overlay experiments reported by Okkels and Anderson (2004) support these findings, with this approach demonstrating interactions between CFP-10/ESAT-6, and Rv0287/Rv0288, but detecting no cross-talk between either of the CFP-10/ESAT-6 and Rv0287/Rv0288 pairs.

To summarise, the yeast two-hybrid studies presented in this thesis strongly suggest that all genome partners within the CFP-10/ESAT-6 family will form stable heterodimeric complexes. However, some members of the family are more promiscuous than others, forming additional complexes with non-genome partners and therefore possibly increasing

the functional flexibility of this protein family. This cross-talk is most likely to occur between closely related family members (groups A/A' and B/B' in figure 5.1), as more distantly related members of the family, such as CFP-10/ESAT-6 (group C/C') and Rv3890c/Rv3891c (group D/D') (Figure 5.1), are almost certainly restricted to interactions with their genome partner. Understanding the rules governing complex formation between the CFP-10/ESAT-6 proteins family provides a useful insight into the potential functional flexibility of this important protein family.

5.1.2 Binding of the CFP-10•ESAT-6 Complex to Specific Host Cells

The fluorescence microscopy assays described in section 3.3.3 provide clear evidence that the CFP-10•ESAT-6 complex specifically binds to the surface of monocyte lineage cells. The results demonstrated the binding of the complex to the surface of primary monocytes/macrophages and monocytic cell lines (U937 and MM6) but not to fibroblast cell lines, which strongly suggests the presence of a specific receptor, or receptors, on the surface of monocytic cells. The fluorescent labelling of the monocytes can be specifically blocked by the addition of excess unlabelled CFP-10•ESAT-6 complex, confirming that the binding is protein specific and is not mediated by the fluorophore.

It has been proposed that the CFP-10•ESAT-6 complex acts as a pore-forming molecule (Hsu *et al.*, 2003), however incubation of the complex with monocyte and macrophage cells for up to 24 hours for the fluorescence microscopy studies provided no evidence of cytolysis. In addition, analysis of surface features of the complex reveals no obvious hydrophobic patch essential for spanning the membrane during pore formation. The surface features also argue against an enzymatic role or binding to nucleic acids. Instead the structural and surface features of the complex, along with the host cell binding studies, are more consistent with a cell signalling role requiring interactions with one or more cell

surface proteins. This proposal is in line with studies by Volkman *et al.* (2004) and Brodin *et al.* (2006) which strongly suggest that the secreted RD1 factors, CFP-10 and ESAT-6, function as key signalling molecules between the mycobacterium and host cell resulting in modulation of host cell activity, such as induction of macrophage aggregation or enhanced phagocytosis of mycobacteria.

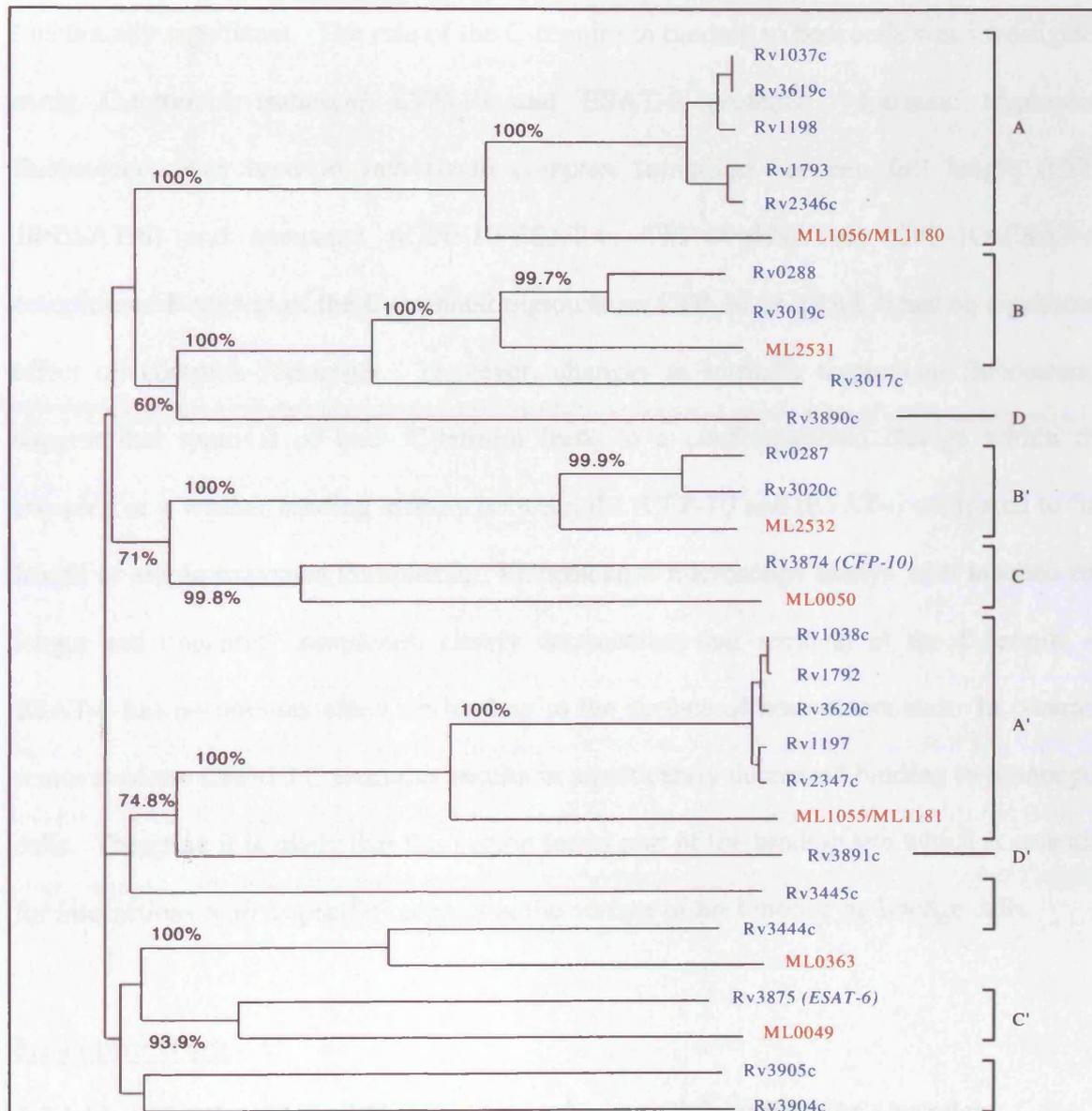


Figure 5.1. Phylogenetic tree for CFP-10/ESAT-6 family of proteins from *M. tuberculosis* (prefixed by Rv, blue) and their *M. leprae* homologues (prefixed by ML, red), prepared by P. Renshaw (Lightbody *et al.*, 2004). Major pairing groups are highlighted by square brackets and labelled (A pairs with A' etc.). Bootstrap values (%) are indicated for the major branch points in the tree.

5.1.3 Host Cell Binding of C-terminal Truncated CFP-10•ESAT-6 Complexes

The most striking features of the CFP-10•ESAT-6 complex are the disordered C-termini of CFP-10 and ESAT-6, which form long flexible arms at either end of the stable four helix bundle. The C-terminal region of CFP-10 and ESAT-6 play no structural role, however the conservation of aromatic and hydrophobic residues suggest that these regions may be functionally significant. The role of the C-termini in binding to host cells was investigated using C-terminal truncated CFP-10 and ESAT-6 proteins. Intrinsic tryptophan fluorescence was used to investigate complex formation between full length (CFP-10•ESAT-6) and truncated (tCFP-10•ESAT-6, CFP-10•tESAT-6, tCFP-10•tESAT-6) complexes. Removal of the C-terminal region from CFP-10 or ESAT-6 has no significant effect on complex formation. However, changes in intrinsic tryptophan fluorescence suggest that removal of both C-termini leads to a conformational change within the complex or a weaker binding affinity between the tCFP-10 and tESAT-6 compared to full length or single truncated complexes. Fluorescence microscopy assays with labelled full length and truncated complexes, clearly demonstrate that removal of the C-termini of ESAT-6 has no obvious effect on binding to the surface of host monocytes. In contrast, removal of the CFP-10 C-terminus results in significantly decreased binding to monocytic cells. Therefore it is likely that this region forms part of the binding site which is essential for interactions with a specific receptor at the surface of host monocyte lineage cells.

5.2 Future Work

5.2.1 Identification of Host Cell Receptors for the CFP-10•ESAT-6 Complex

Following the work described in this thesis, one of the main aims will be to identify the cell surface receptor which interacts with the CFP-10•ESAT-6 complex. Several approaches will be used to identify the CFP-10•ESAT-6 receptor, these include a yeast three-hybrid library screen, photo-activated cross-linking and affinity purification.

Based on the success of the yeast-two hybrid studies investigating complex formation between CFP-10/ESAT-6 family proteins, a yeast three-hybrid approach could be used to identify possible binding partners of the CFP-10•ESAT-6 complex (Zhang and Lautar, 1996, Tirode *et al.*, 1997). The pBridge DNA binding domain (DBD) vector (Clontech) contains two multiple cloning sites (MCS) allowing expression of one protein fused to the GAL4 DBD as well as a second protein of interest. CFP-10 and ESAT-6 have been successfully cloned into the pBridge vector. CFP-10, cloned into MCS 1, is expressed as fusion to the GAL4 DBD and ESAT-6, cloned into MCS 2, is expressed alone. A preliminary screen of the pBridge vector expressing both CFP-10 and ESAT-6 against a human leukocyte cDNA library (Clontech) has resulted in the isolation of over 400 positive colonies. This is a very promising preliminary finding, however, requires a substantial amount of further work to identify the possible binding partners of the CFP-10•ESAT-6 complex.

A recent study involving the cross-linking of peptides from mycobacterial protein Rv2536 to A549 and U937 cells clearly demonstrated an interaction between the peptides and cell membrane proteins (Garcia *et al.*, 2005). A similar approach using the photo-activated cross-linking agent, 4-benzoylbenzoic acid, succinimidyl ester (Molecular Probes), could be used to identify the host cell binding partner, or partners, of the CFP-10•ESAT-6 complex. The N-termini of both CFP-10 and ESAT-6 in the 1:1 complex can be selectively labelled with the succinimidyl ester of 4-benzoylbenzoic acid using the same protocol as described for Alexa Fluor 546 (Molecular Probes). The labelled CFP-10•ESAT-6 complex will be exposed to U937 cells prior to activation of the second reactive group. Upon UV illumination the cross-linking agent will covalently bind the complex to the receptor/receptors, on the surface of host cells. Preliminary studies, involving cross-linking of the CFP-10•ESAT-6 complex are necessary to optimise the

photo-activation process, which is highly dependant on the intensity of the light source and the length of exposure to UV. Following the cross-linking event, the membrane protein fraction from the U937 cells could be separated by SDS-PAGE and bands containing the CFP-10•ESAT-6 complex could be identified by western blotting using CFP-10 or ESAT-6 specific antibodies. N-terminal sequencing or electrospray mass spectrometry of trypsin digested fragments could be used to identify the host cell partner bound to the CFP-10•ESAT-6 complex.

Finally, another option to identify possible binding partners would be affinity purification using a variant of the complex containing an N-terminal His-tag on CFP-10. This allows the CFP-10•ESAT-6 complex to be immobilised on a Ni-NTA (nickel-nitrilotriacetic acid) column, to which the membrane fraction from U937 cells can then be applied. After washing, bound proteins are eluted by imidazole. Using SDS-PAGE analysis eluted fractions are compared to control samples, and proteins detected only in the presence of bound CFP-10•ESAT-6 complex identified by electrospray mass spectroscopy.

5.2.2 Determination of the Function and Mechanism of Action of the CFP-10•ESAT-6 Complex

The function and mechanism of action of a number of very important mycobacterial virulence factors, including the CFP-10•ESAT-6 family complexes and MPT70 (MPB70), remain poorly understood. The data described in this thesis suggests that the CFP-10•ESAT-6 complex functions as a cell signalling molecule, therefore future work will aim to verify this hypothesis, as well as to identify the mechanism of action and the resulting effects on the host cell. Identification of the binding partner, or partners, for the complex would provide essential information on the function of the complex. A potential avenue of investigation may be microarray and proteomics-based studies of the effects of CFP-

10•ESAT-6 on the expression profile of host cells. The studies will aim to identify genes/proteins either up or down-regulated in response to the complex, analysed at various time points over 24 hours.

5.2.3 Investigation of the Structure and Function of CFP-10/ESAT-6 Family Complexes

Like CFP-10 and ESAT-6 other members of the family, including Rv0287, Rv0288 and Rv3019c, have been linked to mycobacterial virulence and pathogenesis (Alderson *et al.*, 2000, Skjot *et al.*, 2000, Skjot *et al.*, 2002, Hogarth *et al.*, 2005). It is thought that all CFP-10/ESAT-6 family complexes will form similar four helix bundles as described for the CFP-10•ESAT-6 complex (Renshaw *et al.*, 2005). Intrinsic tryptophan fluorescence and denaturation studies indicate that Rv0287 and Rv0288 form a 1:1 complex which is significantly more stable than the CFP-10•ESAT-6 complex (Lightbody *et al.*, 2004). NMR (Nuclear Magnetic Resonance) studies will be used to determine the high resolution structure of the Rv0287•Rv0288 complex, and characterise of the surface features of the complex to identify possible functional sites. Following the success of the fluorescence microscopy work described for the CFP-10•ESAT-6 complex, the Rv0287•Rv0288 complex will also be labelled with Alexa Fluor 546 for use in fluorescence microscopy assays to determine whether other CFP-10/ESAT-6 complexes also bind to the surface of monocytic cells.

Appendix

Appendix 1 – Strains, Media and Reagents

1.1 Bacterial and Yeast Strains

E. coli DH5 α - F ϕ 80lacZ Δ M15 Δ (lacZYA-argF)U169 recA1 endA1 hsdR17(r_k⁻, m_k⁺)
phoA supE44 thi-1 gyrA96 relA1 λ ⁻

E. coli BL21 (DE3) - F ompT hsdS_B (r_B⁻ m_B⁻) gal dcm (DE3)

S. cerevisiae CG1935 - MATa, ura3-52, his3-200, ade2-101, lys2-801, trp1-901, leu2-3,
112, gal4-542, gal80 -538, cyhr2, LYS2 : : GAL1_{UAS} -GAL1_{TATA} -HIS3,URA3 : : GAL4₁₇₋
mers(x3) -CYC1_{TATA} -lacZ

1.2 Bacterial Culture Media

Luria-Bertani (LB) Media - 10 g/l tryptone, 5 g/l yeast extract, 10 g/l NaCl. Make up to volume with dH₂O and autoclave. Add appropriate antibiotic when cool. For LB agar add 15 g/l bacteriological agar (Oxoid) prior to autoclaving.

Ampicillin - Ampicillin stock solution: 100 mg/ml in dH₂O and filter sterilised through 0.2 μ m syringe filter. Store at -20 °C. Add to LB media to a final concentration of 100 or 150 μ g/ml as required.

Kanamycin - Stock solution was prepared at 20 mg/ml in dH₂O and filter sterilised through 0.2 μ m syringe filter. Store for one month at -20 °C. Add to LB media to a final concentration of 40 μ g/ml.

1.3 Yeast Culture Media

Yeast Extract Peptone Dextrose (YPD) Media - 20 g/l peptone, 20 g/l glucose, 10 g/l yeast extract. Make up to volume with dH₂O and autoclave. For YPD agar add 15 g/l bacteriological agar (Oxoid) prior to autoclaving.

Synthetic Complete (SC) Drop-out Media - 20 g/l glucose, 6.7 g/l yeast nitrogen base (minus amino acids), 0.6 g/l triple dropout media (-leu/-trp/-his) (BD Biosciences). Make up to 950 ml with dH₂O. Add amino acid supplements as required. Adjust pH to 5.6-5.8, make up to 1 litre with dH₂O and autoclave. For SC drop-out plates add 15 g/l bacteriological agar (Oxoid).

Histidine - Stock solution prepared at 10 mg/ml, use at 20 µg/ml final concentration.

Leucine - Stock solution prepared at 7.2 mg/ml, use at 60 µg/ml final concentration.

Tryptophan - Stock solution prepared at 4.8 mg/ml, use at 40 µg/ml final concentration.

1.4 β-galactosidase Assay Solutions

Z-buffer Solution - 60 mM Na₂HPO₄·7H₂O, 40 mM NaH₂PO₄·H₂O, 10 mM KCl, 1mM MgSO₄ in dH₂O. Store at room temperature.

Z-buffer/X-gal/β-mercaptoethanol Solution - 5 ml Z-buffer, 50 µl X-gal (50 mg/ml in Dimethylformaldehyde - DMF) (Promega), 18 µl β-mercaptoethanol. Store at 4 °C in the dark.

Appendix 2 - Primers

Table of primers used in cloning.

Primer Name	Primer (5' – 3') with underlined restriction sites	Restriction site
CFP-10 Forward	<u>CCG GAA TTC CGG</u> ATG GCA GAG ATG AAG ACC GAT	<i>EcoRI</i>
CFP-10 Reverse	<u>CGC GGA TCC</u> TCA TCA GAA GCC CAT TTG CGA GGA CAG	<i>BamHI</i>
ESAT-6 Forward	<u>CCG GAA TTC CGG</u> ATG ACA GAG CAG CAG TGG AAT	<i>EcoRI</i>
ESAT-6 Reverse	<u>CGC GGA TCC</u> CTA CTA TGC GAA CAT CCC AGT GAC GTT GC	<i>BamHI</i>
Rv0287 Forward	<u>CCG GAA TTC</u> ATG AGC CTT TTG GAT GCT CAT ATC	<i>EcoRI</i>
Rv0287 Reverse	<u>C GGG ATC CCG</u> TCA GAA CCC GGT ATA GGT CGA CGC	<i>BamHI</i>
Rv0288 Forward	<u>CCG GAA TTC</u> ATG TCG CAA ATC ATG TAC AAC TAC	<i>EcoRI</i>
Rv0288 Reverse	<u>C GCG GAT CCT</u> ATG GGT GTT GCA TCC ATG TAG	<i>BamHI</i>
Rv3019c Forward	<u>CCG GAA TTC</u> ATG TCG CAG ATT ATG TAC AAC TAT	<i>EcoRI</i>
Rv3019c Reverse	<u>CGC GGA TCC</u> CTA GCC GCC CCA CTT GGC GGC TTC	<i>BamHI</i>
Rv3020c Forward	<u>CCG GAA TTC</u> ATG AGT TTG TTG TTG GAT GCC CAT ATT	<i>EcoRI</i>
Rv3020c Reverse	<u>CGC GAA TTC</u> TAA AAA CCC GGT GTA GCT GGA CGC	<i>BamHI</i>
tCFP-10 Forward	CAT <u>GCC ATG GCA</u> GAG AGA TGA AGA CCG ATG C	<i>NcoI</i>
tCFP-10 Reverse	CGC <u>GGA TCC</u> TCA GGC CCT CGA GTA TTG GAC GCC	<i>BamHI</i>

Appendix 3 - Protocols

3.1 DNA Analysis by Electrophoresis

Reagents:

Tris-Acetate (TAE) Buffer - 50 x stock solution: 242 g/l Tris Base, 57.1 ml Glacial Acetic Acid, 100 ml 0.5 M EDTA (pH 8.0). Make up to volume with dH₂O.

1 % (w/v) Agarose - 1 g agarose in 100 ml 1 x TAE buffer. Heat until the agarose is fully dissolved. Once cool, add ethidium bromide (12 mg/ml stock solution) to a final concentration of 0.5 µg/ml.

DNA Markers - 100 bp DNA marker and λ DNA *Eco*RI and *Hind*III marker (Promega)

Method:

Prepare samples by adding 5 µl PCR product or plasmid sample, 5 µl dH₂O and 2 µl 6 x Blue/Orange Loading Dye (Promega). Load 12 µl into the wells of a 1 % (w/v) agarose gel. Run at a constant 100 volts for 1 hour in 1 x TAE running buffer. Ethidium bromide is used to visualise the PCR products and plasmids under UV light. Samples for gel extraction are prepared as 20 µl PCR product or plasmid combined with 4 µl 6 x Blue/Orange Loading Dye (Promega).

3.2 Restriction Digest

Reagents:

Sterile dH₂O, 10 x multi-core buffer, 10 µg/ml acetylated BSA and restriction enzymes (*Bam*HI, *Eco*RI and *Nco*I) all supplied by Promega.

Method:

Firstly the appropriate volume of sterile de-ionised H₂O is added to a 1 ml sterile tube. This is followed by the addition of 2 µl 10 x buffer, 0.5 µl 10 µg/ml acetylated BSA and approximately 1 µg DNA. Mix by pipetting. Finally 0.5 µl of each enzyme is added to a final volume of 20 µl. The sample is gently mixed by pipetting and briefly centrifuged in a microcentrifuge before incubating at 37 °C for at least 4 hours.

3.3 Preparation of Competent Cells

Reagents:

LB Medium - 10 ml sterile LB medium with no antibiotic (Appendix 1).

SOB Medium - 2 g tryptone, 0.5 g yeast extract, 0.05 g NaCl. Make up to 95 ml with dH₂O. Add 1 ml 250 mM KCl and adjust pH to 7.0. Make up to 100 ml with dH₂O, divide into 2 x 50 ml in 250 ml flasks and autoclave. Add 0.5 ml sterile 2M MgCl₂ prior to use.

50 mM CaCl₂ Solution - Prepare 50 mM CaCl₂ in dH₂O, filter sterilise through 0.2 µm syringe filter into a sterile container.

50 % Glycerol Solution - Prepare 50 % (v/v) glycerol in dH₂O and sterilise by autoclaving. Store at 4 °C. Use at a final concentration of 15 %.

Method:

Day 1: Streak out *E. coli* (DH5α or BL21 DE3) onto LB agar without antibiotic (Appendix 1). Incubate at 37 °C overnight.

Day 2: Take an individual colony and inoculate into 10 ml sterile LB (minus antibiotic). Incubate overnight at 37 °C in an orbital shaker (200 rpm).

Day 3: Inoculate 50 ml SOB with 1 ml from the overnight culture. Incubate at 37 °C, 200 rpm until the absorbance reads 0.5 at 600 nm. Pellet the cells by centrifugation at 6000 rpm at 4 °C for 15 minutes. Remove the supernatant, resuspend the cell pellet in 25 ml ice-cold 50 mM CaCl₂ and incubate on ice for 30 minutes. Centrifuge cells at 6000 rpm at 4 °C for 15 minutes, remove supernatant and gently resuspend the cell pellet in 2.1 ml ice-cold 50 mM CaCl₂. Add 900 µl ice-cold 50 % glycerol and mix gently by inversion. Aliquot 250 µl into sterile cryotubes then flash freeze in liquid nitrogen. Store at -80 °C.

3.4 Transformation of Competent Cells

Reagents:

LB broth minus antibiotic marker and LB agar plates supplemented with the appropriate antibiotic as described in appendix 1.

Method:

Aliquot 50 µl competent cells into sterile 1 ml tubes. Add 4 µl ligation mix and incubate on ice for 30 minutes. Heat shock at 42 °C for 45 seconds then leave on ice for a further 5 minutes. Add 300 µl LB broth (no antibiotic) to each sample and incubate tubes horizontally at 37 °C at 200 rpm in an orbital shaker for 45 minutes. Plate out 50, 100 and 200 µl aliquots onto LB plates (with appropriate marker) and incubate at 37 °C overnight.

3.5 SDS-Polyacrylamide Gel Electrophoresis

Markers:

Wide Range Molecular Marker – ranging from 6.5 to 205.0 kDa (Sigma)

Ultra Low Range Molecular marker – ranging from 1 to 26.6 kDa (Sigma)

Method:

Add 10 μ l 50 mM DTT and 10 μ l 4 x NuPAGE LDS sample buffer (Invitrogen) to a 20 μ l protein sample. Heat samples at 70 °C for 10 minutes then briefly centrifuge in a microcentrifuge. Load 20 μ l onto a pre-cast NuPAGE Novex 4-12 % Bis-Tris gel (Invitrogen). Run at a constant 200 V, 125 mA and 25 W for 40 minutes in 1 x MES SDS Running Buffer (Invitrogen).

Staining:

Coomassie Brilliant Blue Staining - Dissolve 2.5 g/l coomassie brilliant blue in 40 % methanol (v/v) and 10 % acetic acid (v/v). Stir overnight at room temperature. Make up to 1 litre with dH₂O and filter through 0.2 μ m filter.

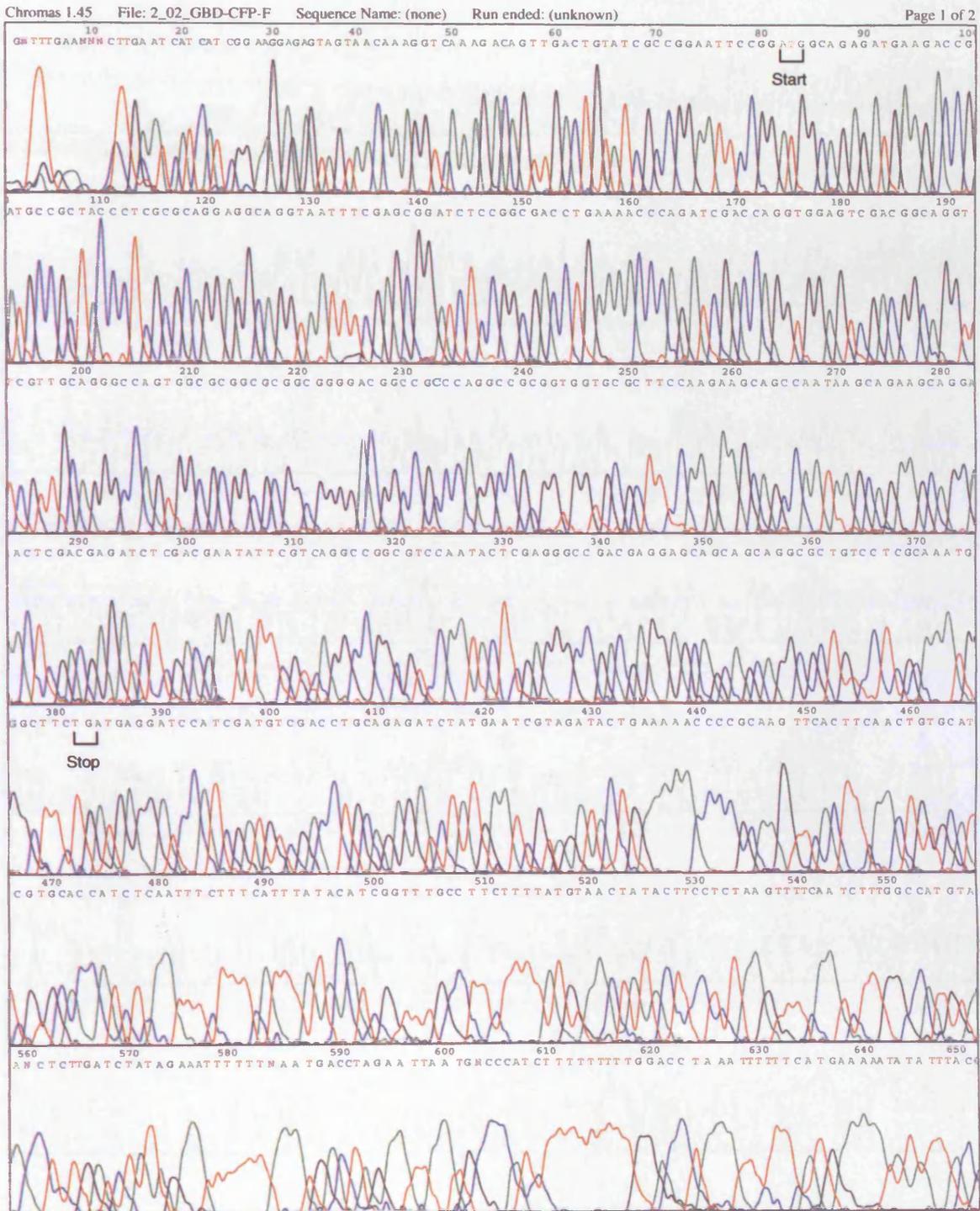
Stain gels in coomassie brilliant blue solution for 20-30 minutes, then destain in 40 % methanol (v/v) and 10 % acetic acid (v/v) solution until protein bands are visible.

Silver staining - Silver Xpress Silver Staining Kit (Invitrogen). Follow the manufacturer's protocol. Briefly, fix the gel in 200 ml fixing solution for 10 minutes. Remove the fixing solution and incubate the gel in two changes of 100 ml sensitising solution, each for 30 minutes. Remove the sensitising solution and wash twice with 200 ml dH₂O for 10 minutes, then incubate the gel in 100 ml staining solution for 15 minutes. Decant the staining solution and rinse with 200 ml dH₂O for 5 minutes. Repeat the wash step before incubating the gel in the 100 ml developing solution, until the bands become visible (3-15 minutes). Once the desired staining intensity has been reached the 5ml of the stopping solution is added and the gel is incubated for a further 10 minutes. The developing

solution/stopping solution is removed and the gel is washed three times in 200 ml dH₂O, each for 10 minutes.

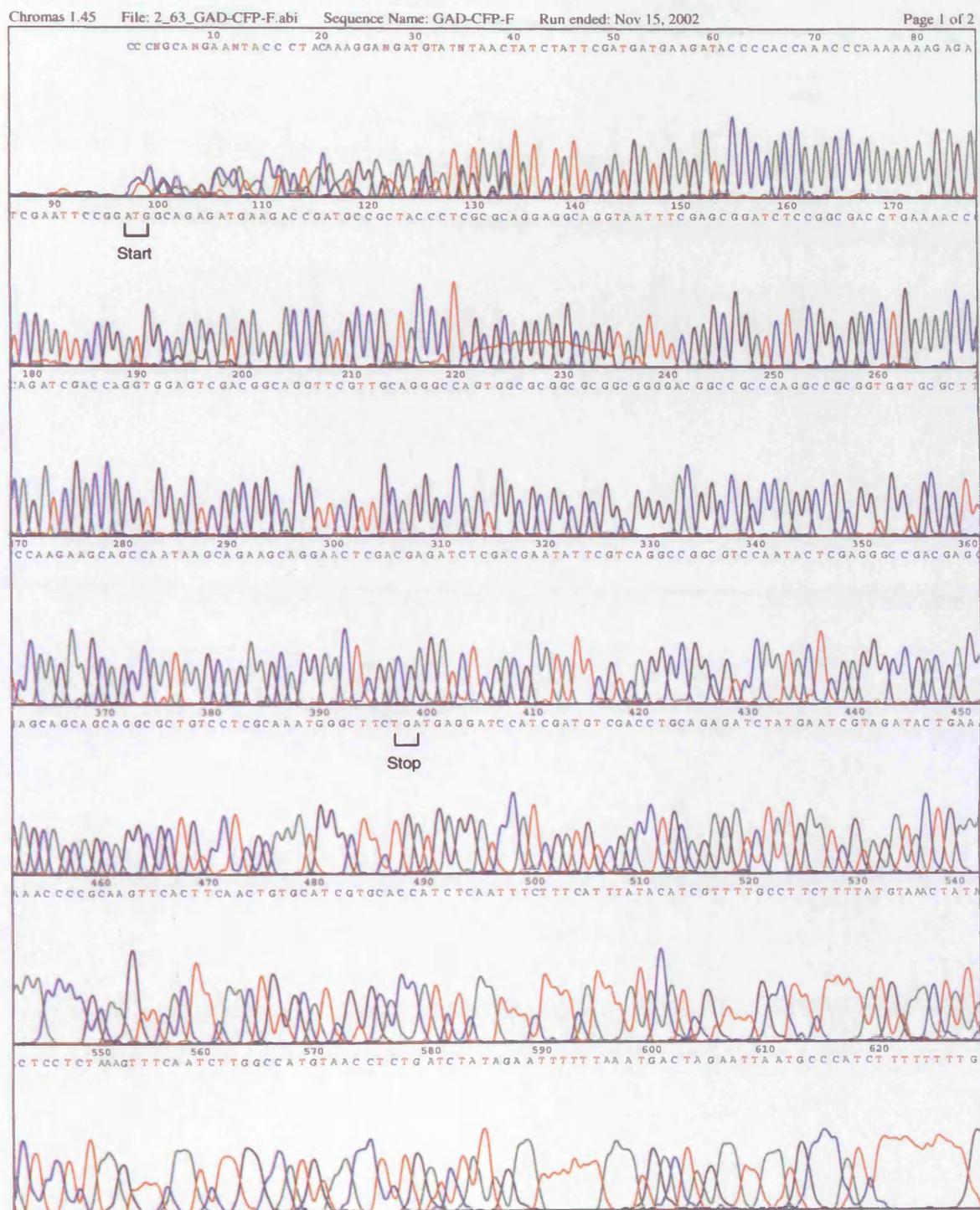
Appendix 4 – DNA Sequencing

4.1 pGBD/CFP-10 Chromatogram



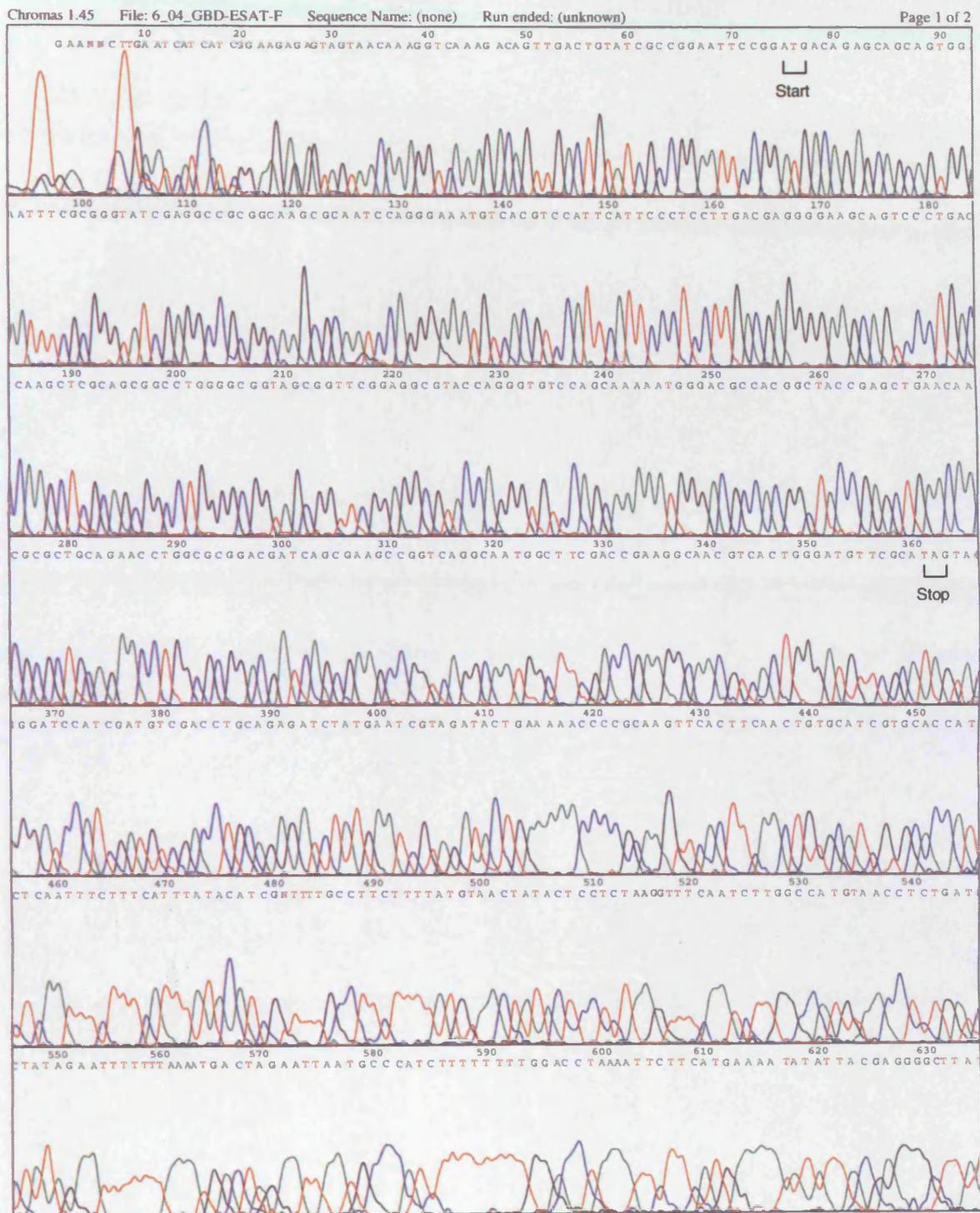
The chromatogram shown here confirms the insertion and integrity of the *cfp-10* gene within the pGBD vector. The *cfp-10* sequence begins from the ATG start codon at position 82 and finishes at the TGA stop codon at position 382.

4.2 pGAD/CFP-10 Chromatogram



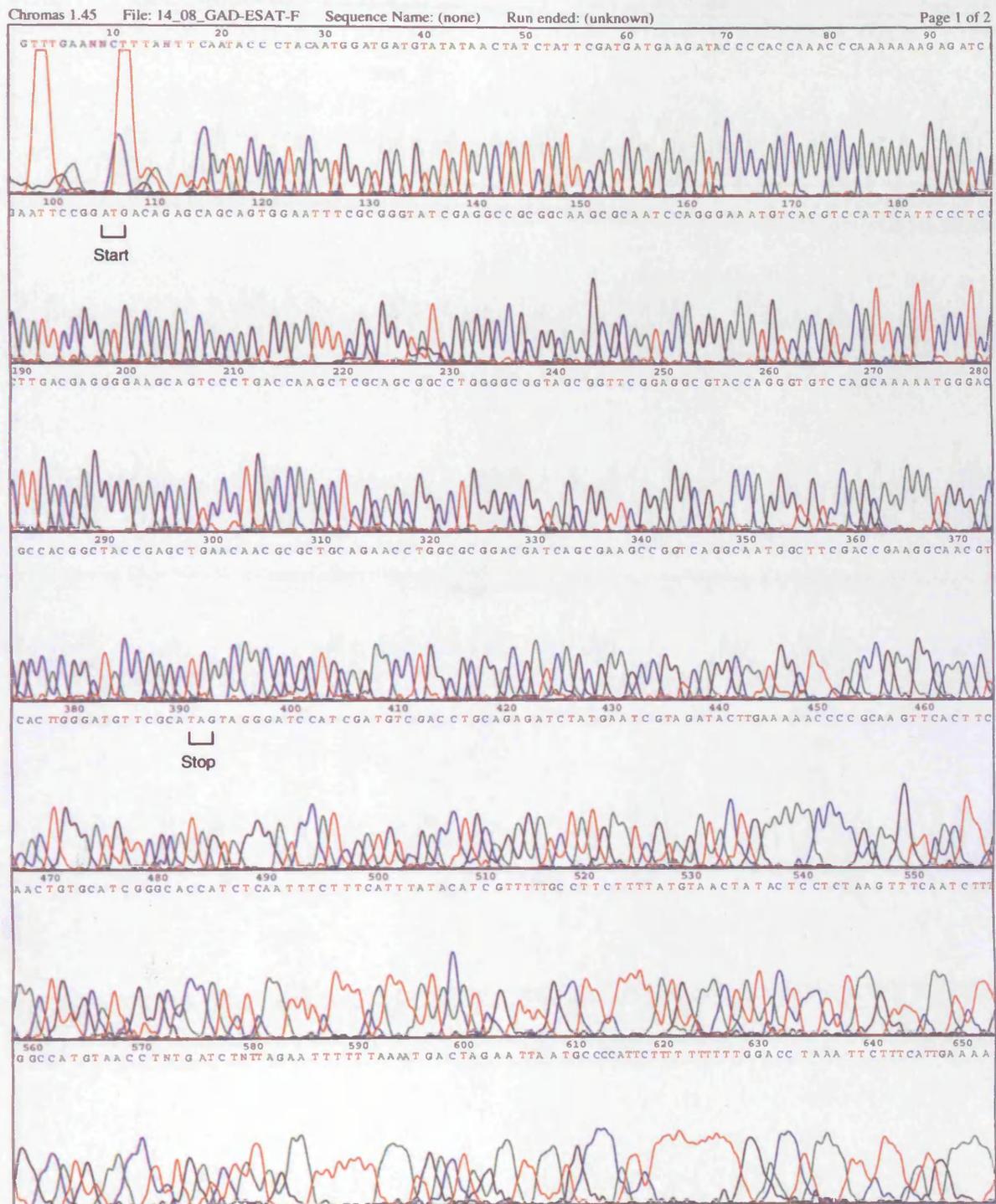
The chromatogram showing the forward sequence of the *cfp-10* gene cloned into the pGAD vector. The *cfp-10* sequence runs from ATG (position 97) to TGA (position 397).

4.3 pGBD/ESAT-6 Chromatogram



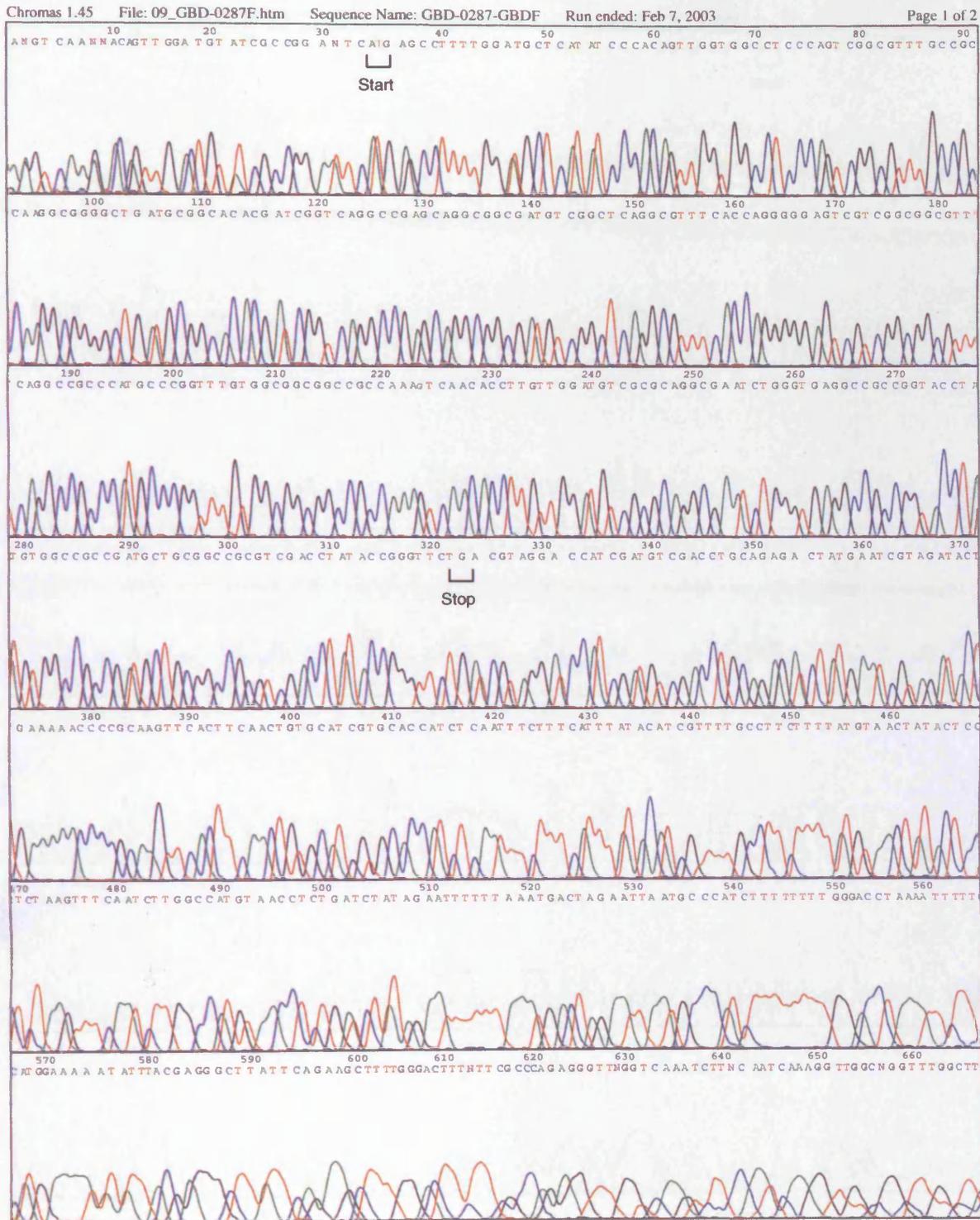
The chromatogram showing insertion of the *esat-6* gene into the pGBD vector. The *esat-6* sequence runs from the ATG start codon (position 75) to the TAG stop codon (position 360).

4.4 pGAD/ESAT-6 Chromatogram



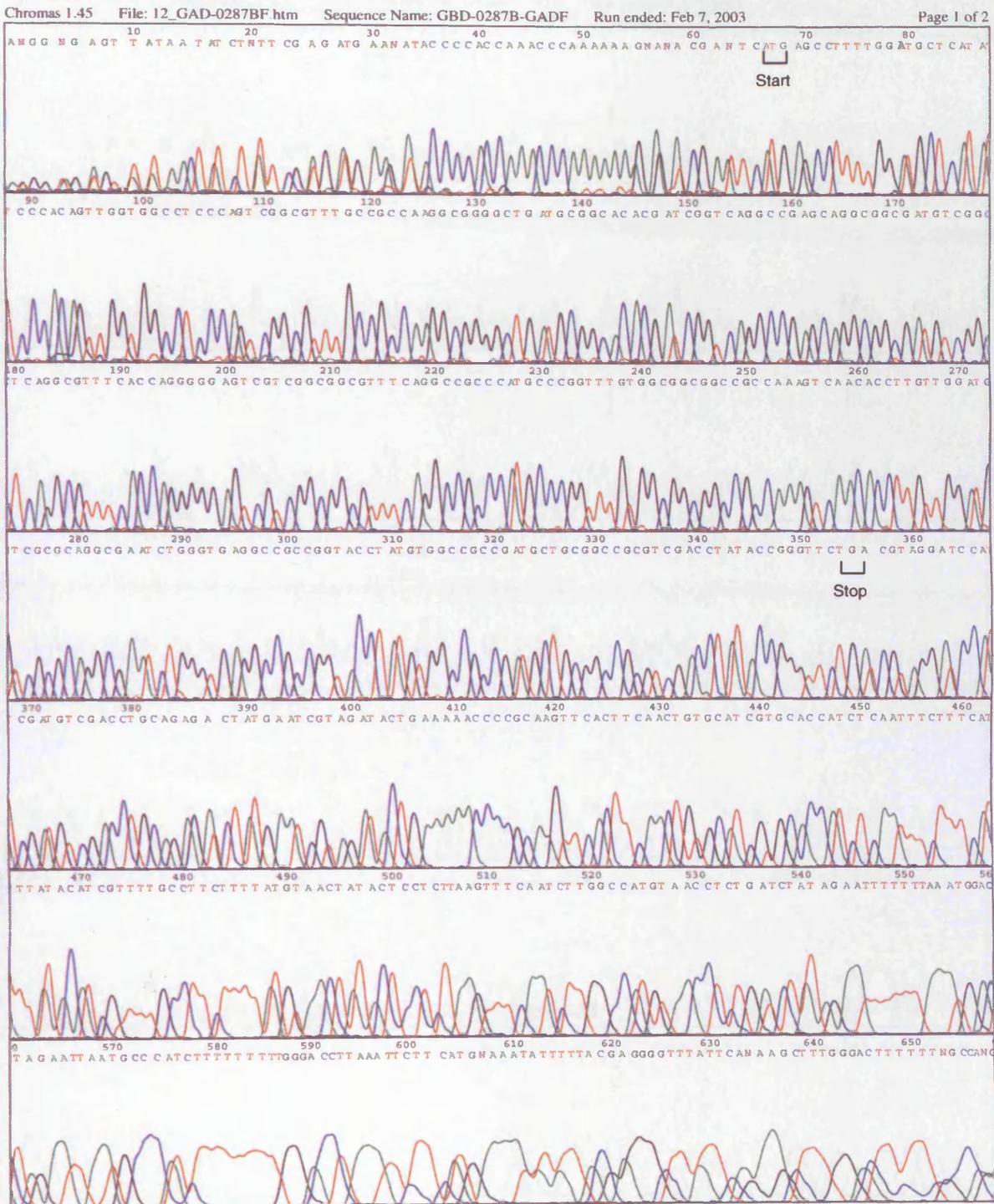
Chromatogram showing the forward sequence of the *esat-6* gene inserted into the pGAD vector. The *esat-6* sequence runs from ATG (position 105) to TAG (position 390).

4.5 pGBD/Rv0287 Chromatogram



The chromatogram shown here confirms the insertion and integrity of the *Rv0287* gene into the pGBD vector. The *Rv0287* sequence runs from the start codon (ATG) at position 34 to the stop codon (TGA) at position 325.

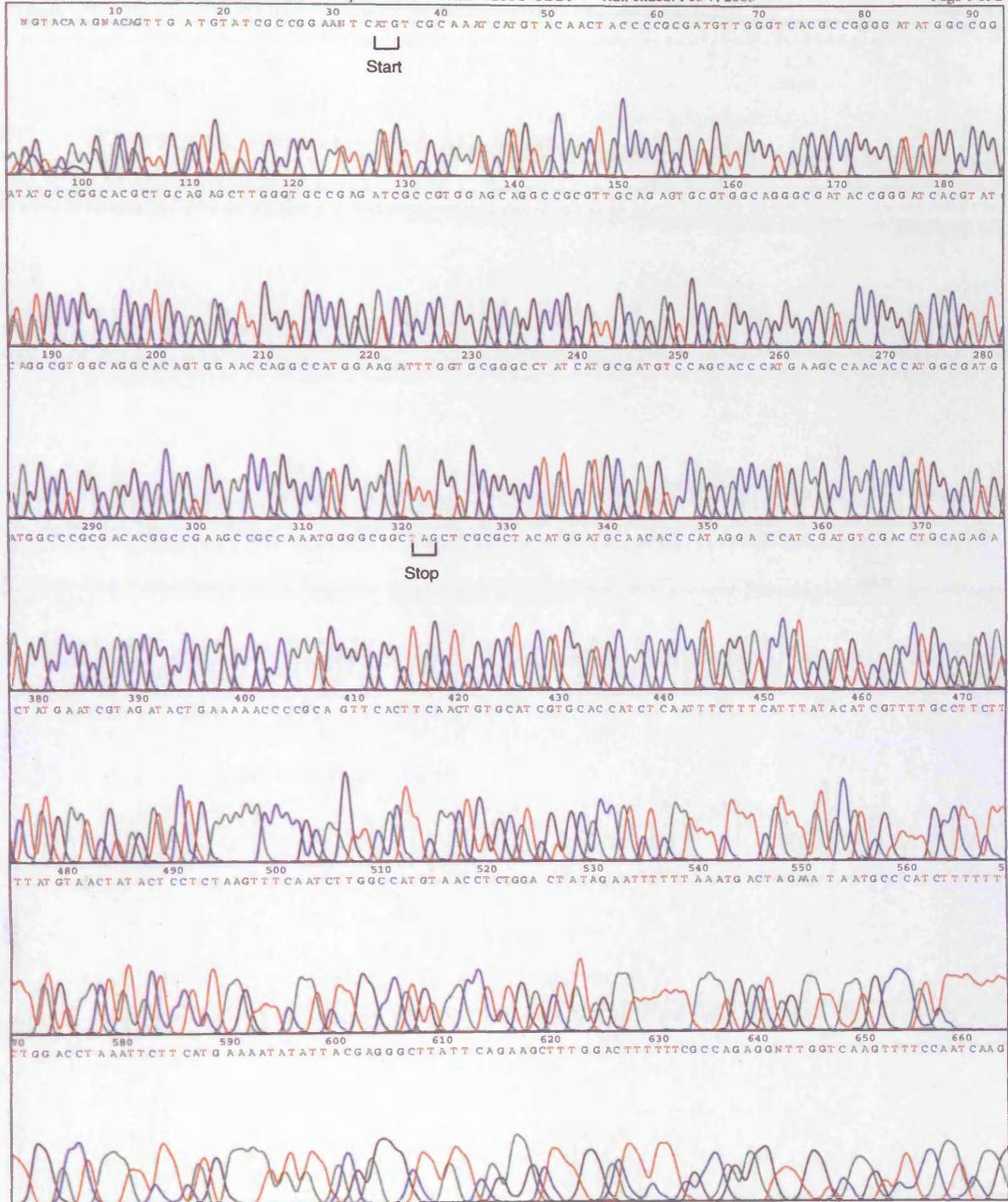
4.6 pGAD/Rv0287 Chromatogram



Chromatogram showing the forward sequence of the *Rv0287* gene inserted into the pGAD vector. The *Rv0287* sequence runs from ATG (position 73) to TGA (position 361).

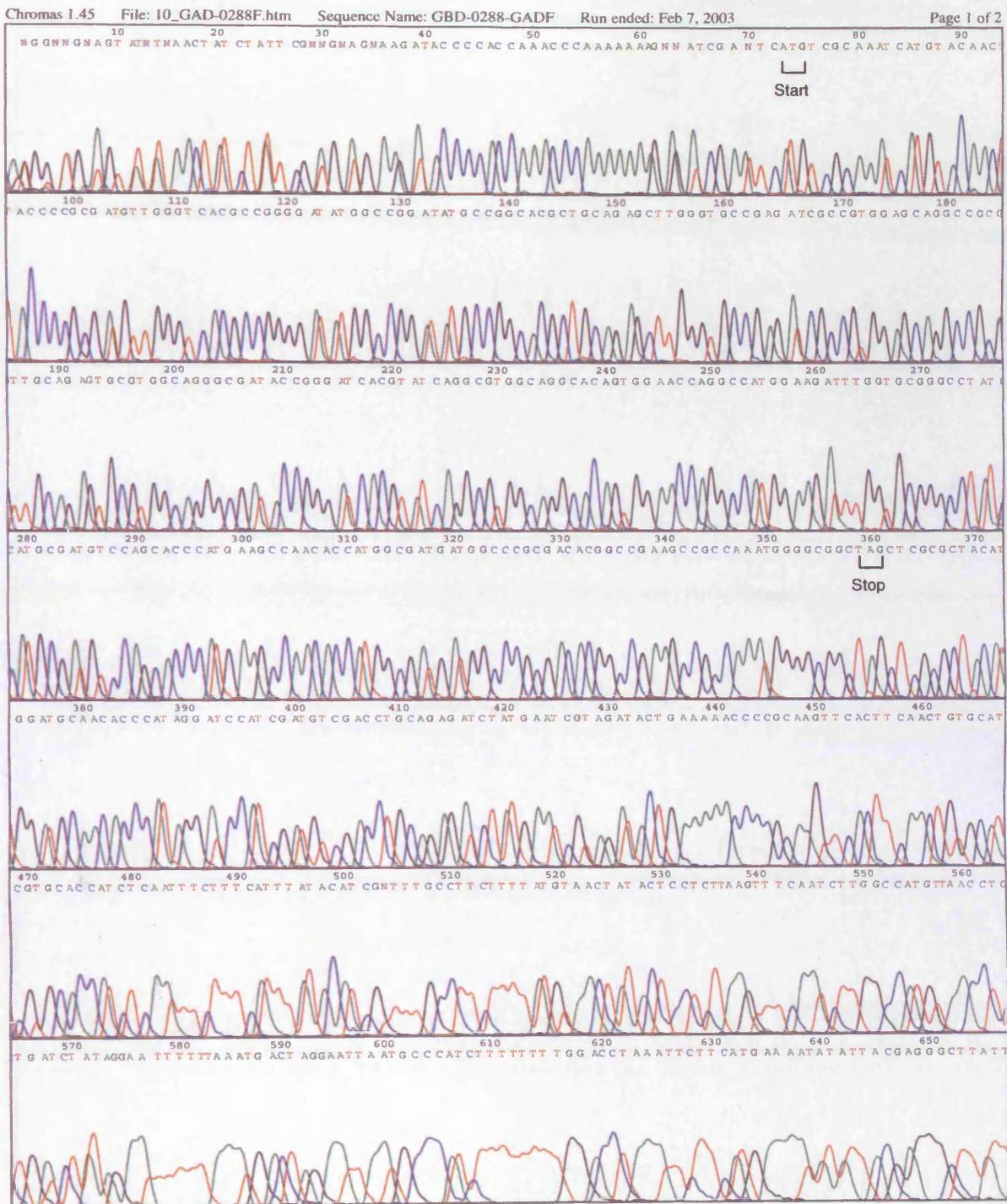
4.7 pGBD/Rv0288Chromatogram

Chromas 1.45 File: 17_GBD-0288CF.htm Sequence Name: GBD-0288C-GBDF Run ended: Feb 7, 2003 Page 1 of 2



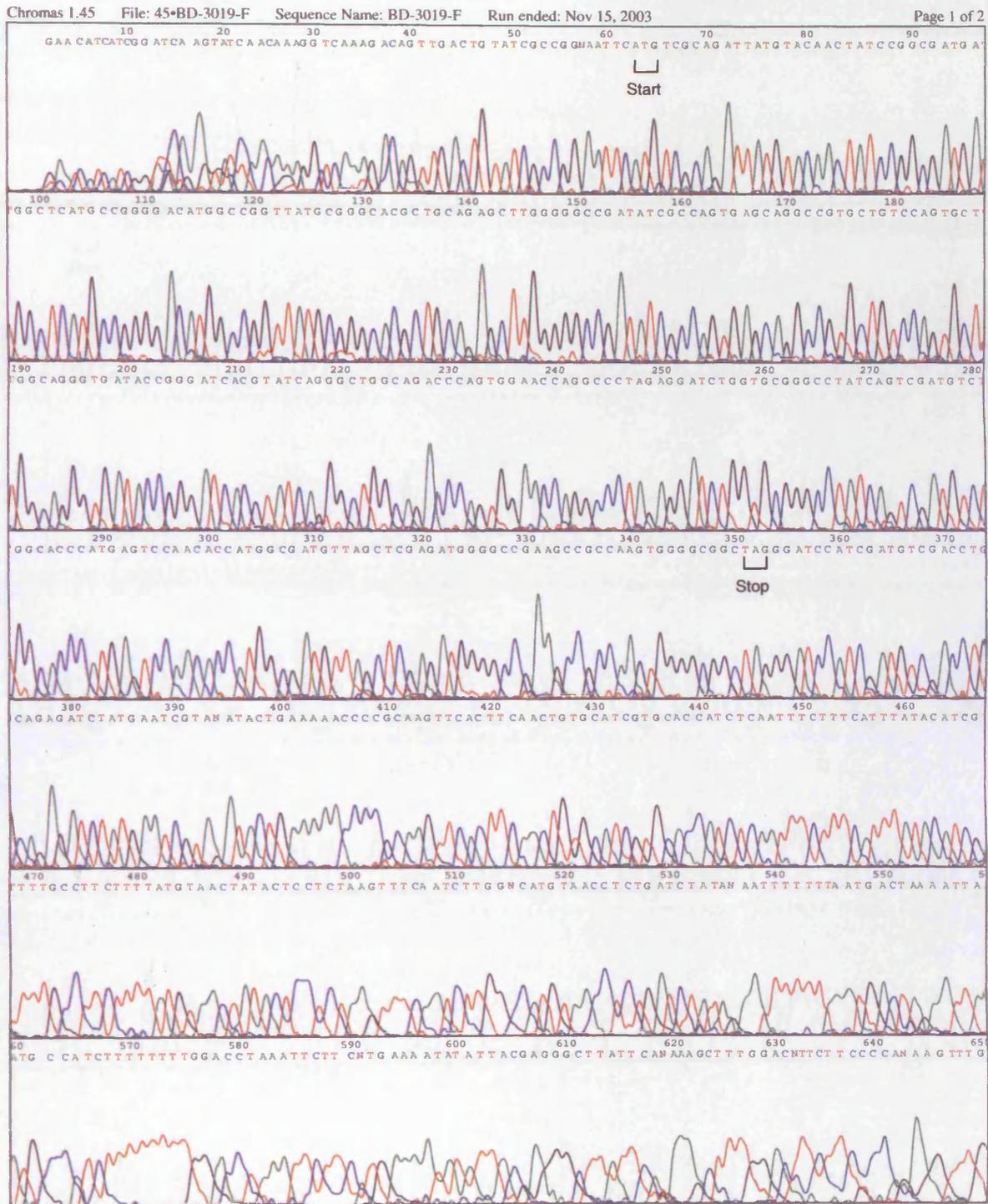
Chromatogram showing the forward sequence of the *Rv0288* gene inserted into the pGBD vector. The *Rv0288* sequence runs from ATG (position 35) to TAG (position 323).

4.8 pGAD/Rv0288 Chromatogram



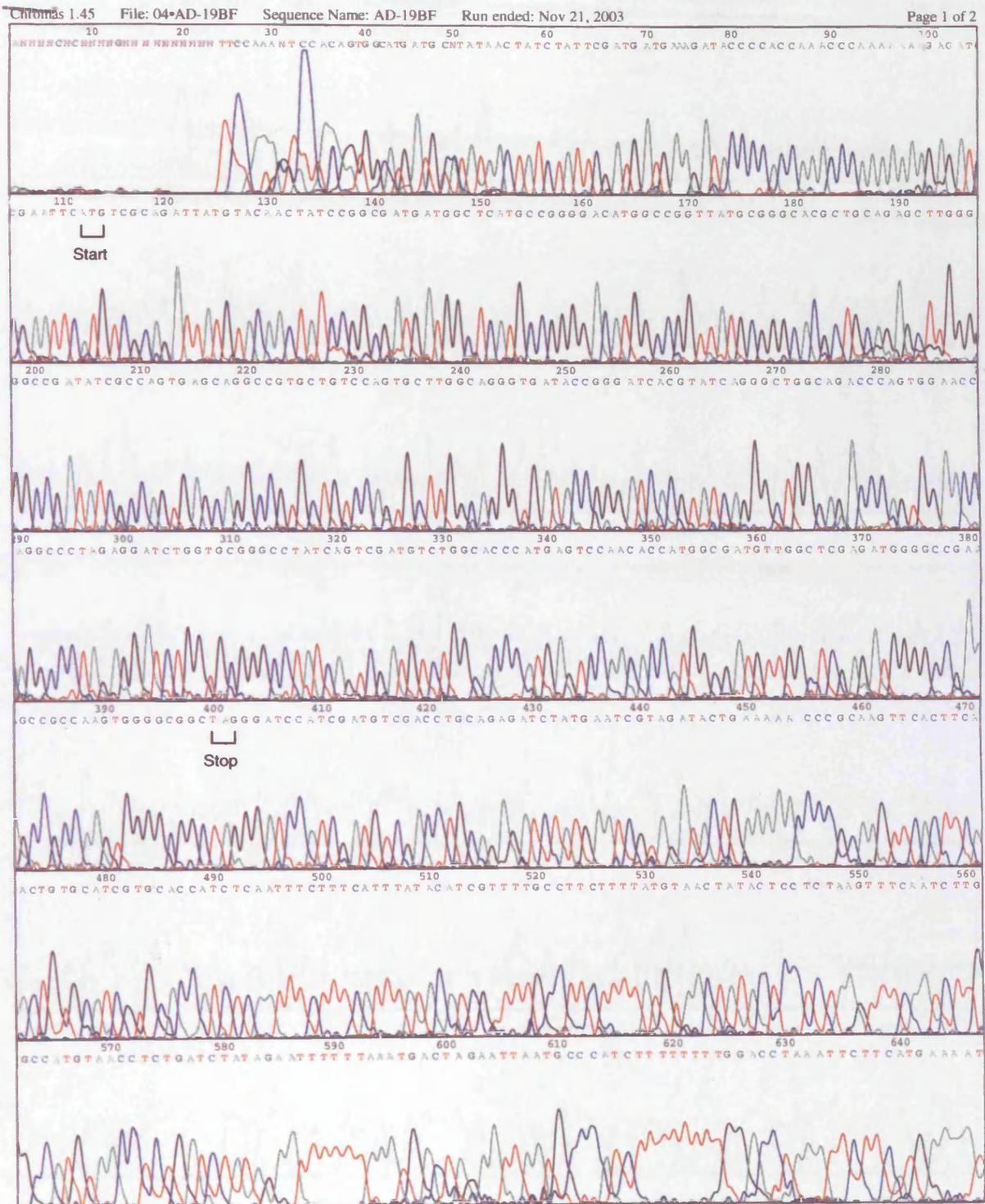
Chromatogram showing the integrity and insertion of the *Rv0288* gene into the pGAD vector. The *Rv0288* sequence runs from the ATG start codon (position 73) to the TAG stop codon (position 361).

4.9 pGBD/Rv3019c Chromatogram



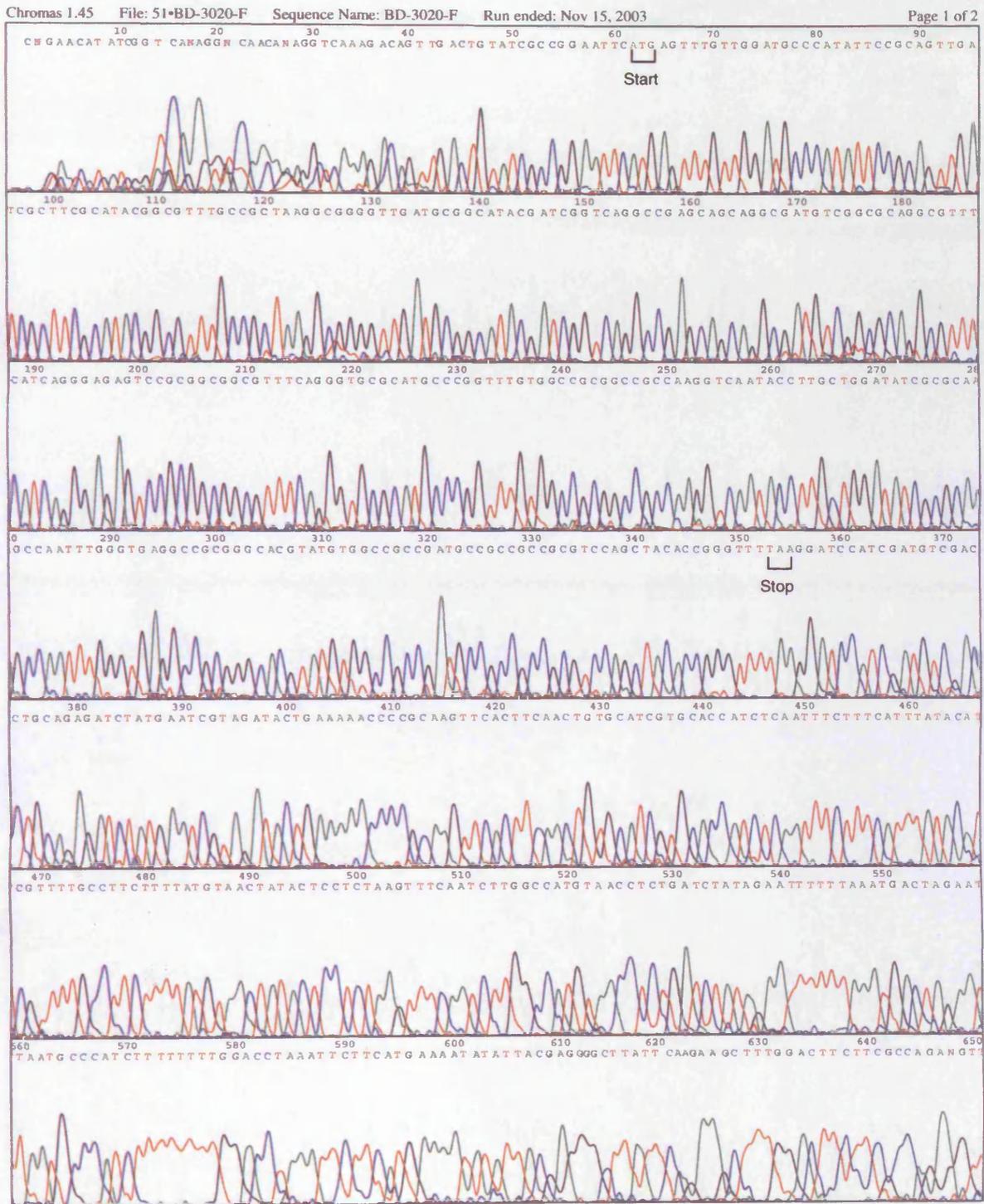
Chromatogram showing insertion of the *Rv3019c* gene into the pGBD vector. The *Rv3019c* sequence runs from the ATG start codon at position 63 to the TAG stop codon at position 354. There is a silent mutation at position 314 where a G is altered to an A, however both TTA and TTG encode for leucine.

4.10 pGAD/Rv3019c Chromatogram



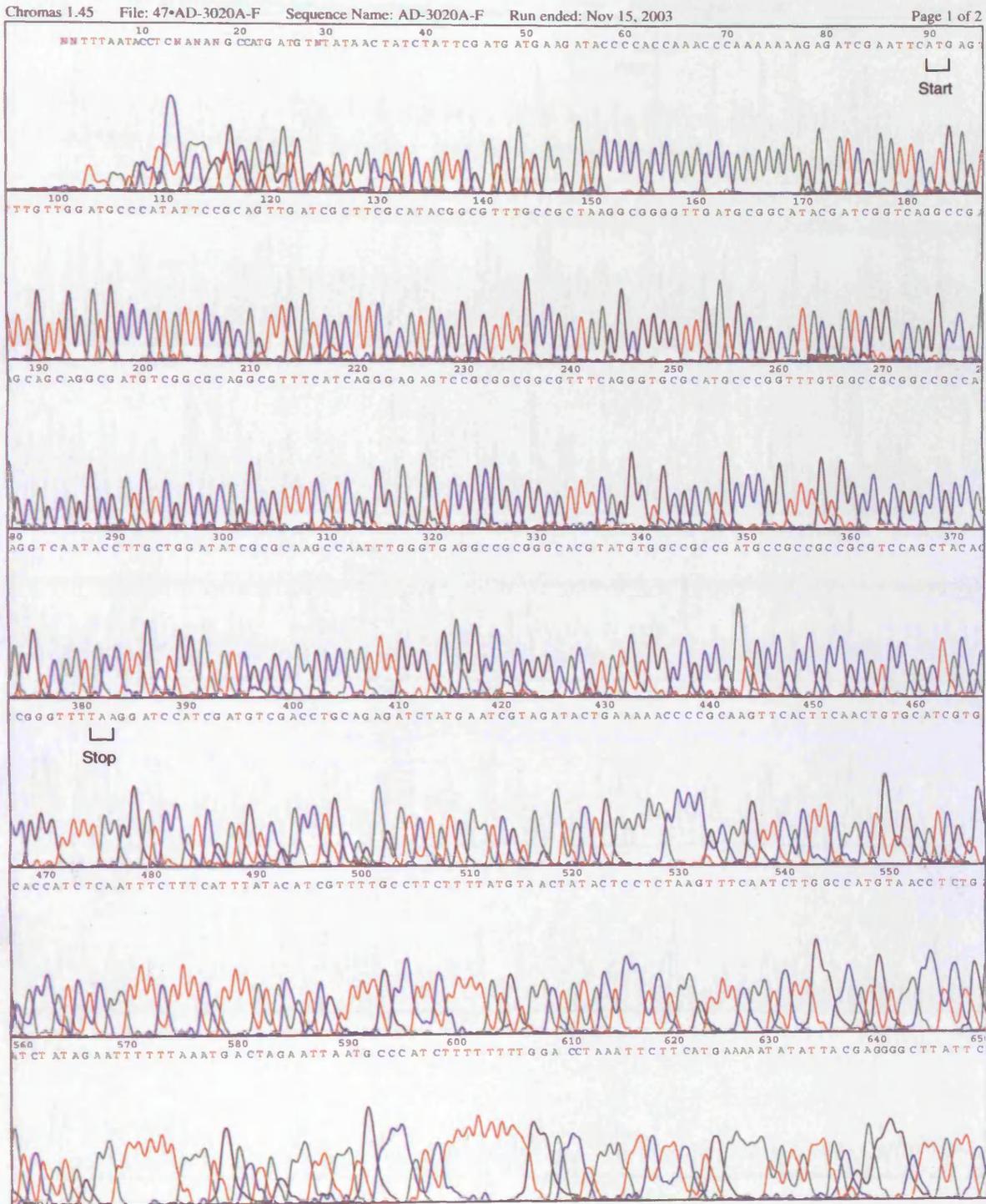
Chromatogram confirming insertion of the *Rv3019c* gene into the pGAD vector. The *Rv3019c* sequence runs from the ATG start codon at position 112 to the TAG stop codon at position 400.

4.11 pGBD/Rv3020c Chromatogram



Chromatogram confirming the integrity and insertion of the *Rv3020c* gene into the pGBD vector. The *Rv3020c* sequence runs from ATG (position 63) to TAA (position 353).

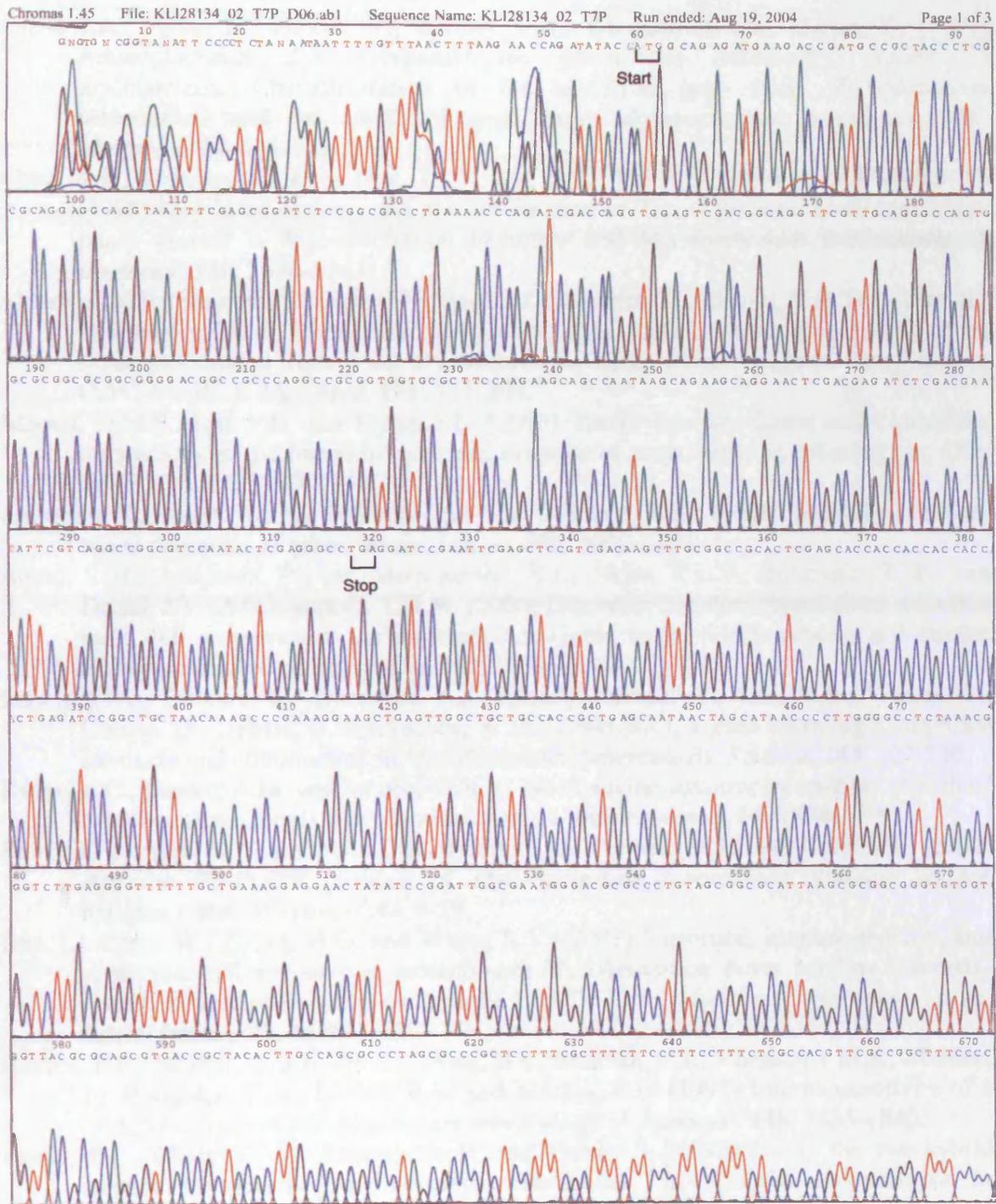
4.12 pGAD/Rv3020c Chromatogram



Chromatogram showing the forward sequence of the *Rv3020c* gene in the pGAD vector.

The *Rv3020c* sequence runs from the ATG start codon at position 90 to the TAA stop codon at position 381.

4.13 pET28a/tCFP-10 Chromatogram



The chromatogram shown above confirms the integrity and insertion of the truncated *cfp-10* sequence in the pET28a vector. The truncated *cfp-10* sequence runs from the ATG start codon at position 60 to the TGA stop codon at position 218. The tCFP-10 sequence is 258 bp in length, lacking the C-terminal 42 residues (12 amino acids).

References

- Ainsa, J.A., Perez, E., Pelicic, V., Berthet, F.X., Gicquel, B. and Martin, C. (1997) Aminoglycoside 2'-N-acetyltransferase genes are universally present in mycobacteria: Characterization of the *aac(2')*-Ic gene from *Mycobacterium tuberculosis* and the *aac(2')*-Id gene from *Mycobacterium smegmatis*. *Mol. Microbiol.* **24**, 431-441.
- Ainsa, J.A., Blokpoel, M.C.J., Ota, I., Young, D.B., De Smet, K.A.L. and Martin, C. (1998) Molecular cloning and characterization of Tap, a putative multidrug efflux pump present in *Mycobacterium fortuitum* and *Mycobacterium tuberculosis*. *J. Bacteriol.* **180**, 5836-5843.
- Alderson, M.R., Bement, T., Day, C.H., Zhu, L.Q., Molesh, D., Skeiky, Y.A.W., Coler, R., Lewinsohn, D.M., Reed, S.G. and Dillon, D.C. (2000) Expression cloning of an immunodominant family of *Mycobacterium tuberculosis* antigens using human CD4⁺ T cells. *J. Exp. Med.* **191**, 551-559.
- Algood, H.M.S., Lin, P.L. and Flynn, J.L. (2005) Tumor necrosis factor and chemokine interactions in the formation and maintenance of granulomas in tuberculosis. *Clin. Infect. Dis.* **41**, S189-S193.
- Andersen, P., Munk, M.E., Pollock, J.M. and Doherty, T.M. (2000) Specific immune-based diagnosis of tuberculosis. *Lancet.* **356**, 1099-1104.
- Arend, S.M., Andersen, P., van Meijgaarden, K.E., Skjot, R.L.V., Subronto, Y.W., van Dissel, J.T. and Ottenhoff, T.H.M. (2000) Detection of active tuberculosis infection by T cell responses to early-secreted antigenic target 6-kDa protein and culture filtrate protein 10. *J. Infect. Dis.* **181**, 1850-1854.
- Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K.S., Wilson, T., Collins, D., Delisle, G. and Jacobs, W.R. (1994) *InhA*, a gene encoding a target for isoniazid and ethionamide in *Mycobacterium tuberculosis*. *Science.* **263**, 227-230.
- Bange, F.C., Brown, A.M. and Jacobs, W.R. (1996) Leucine auxotrophy restricts growth of *Mycobacterium bovis* BCG in macrophages. *Infect. Immun.* **64**, 1794-1799.
- Banu, S., Honore, N., Saint-Joanis, B., Philpott, D., Prevost, M.C. and Cole, S.T. (2002) Are the PE-PGRS proteins of *Mycobacterium tuberculosis* variable surface antigens? *Mol. Microbiol.* **44**, 9-19.
- Bao, L., Chen, W., Zhang, H.D. and Wang, X.Y. (2003) Virulence, immunogenicity, and protective efficacy of two recombinant *Mycobacterium bovis* bacillus Calmette-Guerin strains expressing the antigen ESAT-6 from *Mycobacterium tuberculosis*. *Infect. Immun.* **71**, 1656-1661.
- Barnes, P.F., Mehra, V., Rivoire, B., Fong, S.J., Brennan, P.J., Voegtline, M.S., Minden, P., Houghten, R.A., Bloom, B.R. and Modlin, R.L. (1992) Immunoreactivity of a 10-kDa antigen of *Mycobacterium tuberculosis*. *J. Immunol.* **148**, 1835-1840.
- Bartel, P.L., Chien, C. T., Strenglanz, R. and Fields, S. (1993) Using the two-hybrid system to detect protein-protein interactions. In *Cellular Interactions in Development: A Practical Approach*. pp. 153-179.
- Bean, A.G.D., Roach, D.R., Briscoe, H., France, M.P., Korner, H., Sedgwick, J.D. and Britton, W.J. (1999) Structural deficiencies in granuloma formation in TNF gene targeted mice underlie the heightened susceptibility to aerosol *Mycobacterium tuberculosis* infection, which is not compensated for by lymphotoxin. *J. Immunol.* **162**, 3504-3511.
- Behr, M.A., Wilson, M.A., Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S. and Small, P.M. (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science.* **284**, 1520-1523.
- Behr, M.A. and Small, P.M. (1997) Has BCG attenuated to impotence? *Nature.* **389**, 133-134.

- Berthet, F.X., Rasmussen, P.B., Rosenkrands, I., Andersen, P. and Gicquel, B. (1998) A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). *Microbiology*. **144**, 3195-3203.
- Betts, M.J. and Russell, R.B. (2003) Amino acid properties and consequences of substitutions. In 'Bioinformatics for Geneticists'. Eds. Barnes, M.R. and Gray, I.C. John Wiley & Sons, Chichester.
- Betz, S.F., Liebman, P.A. and DeGrado, W.F. (1997) De novo design of native proteins: Characterization of proteins intended to fold into antiparallel, rop-like, four-helix bundles. *Biochemistry*. **36**, 2450-2458.
- Bloemink, M.J., Kemmink, J., Dentten, E., Muskett, F.W., Whelan, A., Sheikh, A., Hewinson, G., Williamson, R.A. and Carr, M.D. (2001) Sequence-specific assignment and determination of the secondary structure of the 163-residue *M. tuberculosis* and *M. bovis* antigenic protein MPB70. *J. Biomol. NMR*. **20**, 185-186.
- Brandt, L., Cunha, J.F., Olsen, A.W., Chilima, B., Hirsch, P., Appelberg, R. and Andersen, P. (2002) Failure of the *Mycobacterium bovis* BCG vaccine: Some species of environmental mycobacteria block multiplication of BCG and induction of protective immunity to tuberculosis. *Infect. Immun.* **70**, 672-678.
- Breeden, L. and Nasmyth, K. (1985) Regulation of the yeast *HO* gene. *Cold Spring Harbor Symposia on Quantitative Biology*. **50**, 643-650.
- Brennan, M.J., Delogu, G., Chen, Y.P., Bardarov, S., Kriakov, J., Alavi, M. and Jacobs, W.R. (2001) Evidence that mycobacterial PE_PGRS proteins are cell surface constituents that influence interactions with other cells. *Infect. Immun.* **69**, 7326-7333.
- Brent, R. and Ptashne, M. (1985) A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell*. **43**, 729-736.
- Brodin, P., Eiglmeier, K., Marmiesse, M., Billault, A., Garnier, T., Niemann, S., Cole, S.T. and Brosch, R. (2002) Bacterial artificial chromosome-based comparative genomic analysis identifies *Mycobacterium microti* as a natural ESAT-6 deletion mutant. *Infect. Immun.* **70**, 5568-5578.
- Brodin, P., de Jonge, M.I., Majlessi, L., Leclerc, C., Nilges, M., Cole, S.T. and Brosch, R. (2005) Functional analysis of early secreted antigenic target-6, the dominant T-cell antigen of *Mycobacterium tuberculosis*, reveals key residues involved in secretion, complex formation, virulence, and immunogenicity. *J. Biol. Chem.* **280**, 33953-33959.
- Brodin, P., Majlessi, L., Marsollier, L., de Jonge, M.I., Bottai, D., Demangel, C., Hinds, J., Neyrolles, O., Butcher, P.D., Leclerc, C., Cole, S.T. and Brosch, R. (2006) Dissection of ESAT-6 system 1 of *Mycobacterium tuberculosis* and impact on immunogenicity and virulence. *Infect. Immun.* **74**, 88-98.
- Brosch, R., Gordon, S.V., Billault, A., Garnier, T., Eiglmeier, K., Soravito, C., Barrell, B.G. and Cole, S.T. (1998) Use of a *Mycobacterium tuberculosis* H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. *Infect. Immun.* **66**, 2221-2229.
- Brosch, R., Pym, A.S., Gordon, S.V. and Cole, S.T. (2001) The evolution of mycobacterial pathogenicity: clues from comparative genomics. *Trends Microbiol.* **9**, 452-458.
- Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D. and Cole, S.T. (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 3684-3689.
- Brusasca, P.N., Colangeli, R., Lyashchenko, K.P., Zhao, X., Vogelstein, M., Spencer, J.S., McMurray, D.N. and Gennaro, M.L. (2001) Immunological characterization of

antigens encoded by the RD1 region of the *Mycobacterium tuberculosis* genome. *Scand. J. Immunol.* **54**, 448-452.

- Buddle, B.M., Parlane, N.A., Keen, D.L., Aldwell, F.E., Pollock, J.M., Lightbody, K. and Andersen, P. (1999) Differentiation between *Mycobacterium bovis* BCG-vaccinated and *M. bovis* infected cattle by using recombinant mycobacterial antigens. *Clin. Diagn. Lab. Immunol.* **6**, 1-5.
- Burke, B., Giannoudis, A., Corke, K.P., Gill, D., Wells, M., Ziegler-Heitbrock, L. and Lewis, C.E. (2003) Hypoxia-induced gene expression in human macrophages: Implications for ischemic tissues and hypoxia-regulated gene therapy. *Am. J. Pathol.* **163**, 1233-1243.
- Burts, M.L., Williams, W.A., DeBord, K. and Missiakas, D.M. (2005) EsxA and EsxB are secreted by an ESAT-6-like system that is required for the pathogenesis of *Staphylococcus aureus* infections. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 1169-1174.
- Camus, J.C., Pryor, M.J., Medigue, C. and Cole, S.T. (2002) Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology.* **148**, 2967-2973.
- Carr, M.D., Bloemink, M.J., Dentten, E., Whelan, A.O., Gordon, S.V., Kelly, G., Frenkiel, T.A., Hewinson, R.G. and Williamson, R.A. (2003) Solution structure of the *Mycobacterium tuberculosis* complex protein MPB70: From tuberculosis pathogenesis to inherited human corneal disease. *J. Biol. Chem.* **278**, 43736-43743.
- Chambers, M.A., Williams, A., Gavier-Widen, D., Whelan, A., Hall, G., Marsh, P.D., Bloom, B.R., Jacobs, W.R. and Hewinson, R.G. (2000) Identification of a *Mycobacterium bovis* BCG auxotrophic mutant that protects guinea pigs against *M. bovis* and hematogenous spread of *Mycobacterium tuberculosis* without sensitization to tuberculin. *Infect. Immun.* **68**, 7094-7099.
- Chambers, M.A., Williams, A., Hatch, G., Gavier-Widen, D., Hall, G., Huygen, K., Lowrie, D., Marsh, P.D. and Hewinson, R.G. (2002) Vaccination of guinea pigs with DNA encoding the mycobacterial antigen MPB83 influences pulmonary pathology but not hematogenous spread following aerogenic infection *Mycobacterium bovis*. *Infect. Immun.* **70**, 2159-2165.
- Charlet, D., Mostowy, S., Alexander, D., Sit, L., Wiker, H.G. and Behr, M.A. (2005) Reduced expression of antigenic proteins MPB70 and MPB83 in *Mycobacterium bovis* BCG strains due to a start codon mutation in sigK. *Mol. Microbiol.* **56**, 1302-1313.
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497-3500.
- Chien, C.T., Bartel, P.L., Sternglanz, R. and Fields, S. (1991) The two-hybrid system: A method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Nat. Acad. Sci. U.S.A.* **88**, 9578-9582.
- Choudhuri, B.S., Bhakta, S., Barik, R., Basu, J., Kundu, M. and Chakrabarti, P. (2002) Overexpression and functional characterization of an ABC (ATP-binding cassette) transporter encoded by the genes *drxA* and *drxB* of *Mycobacterium tuberculosis*. *Biochem. J.* **367**, 279-285.
- Clemens, D.L. and Horwitz, M.A. (1995) Characterization of the *Mycobacterium tuberculosis* phagosome and evidence that phagosomal maturation is inhibited. *J. Exp. Med.* **181**, 257-270.
- Clemens, D.L. and Horwitz, M.A. (1996) The *Mycobacterium tuberculosis* phagosome interacts with early endosomes and is accessible to exogenously administered transferrin. *J. Exp. Med.* **184**, 1349-1355.
- Cockle, P.J., Gordon, S.V., Lalvani, A., Buddle, B.M., Hewinson, R.G. and Vordermeier, H.M. (2002) Identification of novel *Mycobacterium tuberculosis* antigens with

- potential as diagnostic reagents or subunit vaccine candidates by comparative genomics. *Infect. Immun.* **70**, 6996-7003.
- Colangeli, R., Spencer, J.S., Bifani, P., Williams, A., Lyashchenko, K., Keen, M.A., Hill, P.J., Belisle, J. and Gennaro, M.L. (2000) MTSA-10, the product of the Rv3874 gene of *Mycobacterium tuberculosis*, elicits tuberculosis-specific, delayed-type hypersensitivity in guinea pigs. *Infect. Immun.* **68**, 990-993.
- Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S.V., Eiglmeier, K., Gas, S., Barry, C.E., Tekaiia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M.A., Rajandream, M.A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Taylor, K., Whitehead, S. and Barrell, B.G. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature.* **393**, 537-544.
- Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D., Thomson, N.R., Wheeler, P.R., Honore, N., Garnier, T., Churcher, C., Harris, D., Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R.M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C., Maclean, J., Moule, S., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S., Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K., Taylor, K., Whitehead, S., Woodward, J.R. and Barrell, B.G. (2001) Massive gene decay in the leprosy bacillus. *Nature.* **409**, 1007-1011.
- Collins, D.M., Kawakami, R.P., Wards, B.J., Campbell, S. and de Lisle, G.W. (2003) Vaccine and skin testing properties of two avirulent *Mycobacterium bovis* mutants with and without an additional esat-6 mutation. *Tuberculosis.* **83**, 361-366.
- Converse, S.E. and Cox, J.S. (2005) A protein secretion pathway critical for *Mycobacterium tuberculosis* virulence is conserved and functional in *Mycobacterium smegmatis*. *J. Bacteriol.* **187**, 1238-1245.
- De la Rúa-Domenech, R. (2005) Human *Mycobacterium bovis* infection in the United Kingdom: Incidence, risks, control measures and review of the zoonotic aspects of bovine tuberculosis. *Tuberculosis.* In Press.
- Delogu, G. and Brennan, M.J. (2001) Comparative immune response to PE and PE_PGRS antigens of *Mycobacterium tuberculosis*. *Infect. Immun.* **69**, 5606-5611.
- Demangel, C., Brodin, P., Cockle, P.J., Brosch, R., Majlessi, L., Leclerc, C. and Cole, S.T. (2004) Cell envelope protein PPE68 contributes to *Mycobacterium tuberculosis* RD1 immunogenicity independently of a 10-kilodalton culture filtrate protein and ESAT-6. *Infect. Immun.* **72**, 2170-2176.
- Derrick, S.C., Repique, C., Snoy, P., Yang, A.L. and Morris, S. (2004a) Immunization with a DNA vaccine cocktail protects mice lacking CD4 cells against an aerogenic infection with *Mycobacterium tuberculosis*. *Infect. Immun.* **72**, 1685-1692.
- Derrick, S.C., Yang, A.L. and Morris, S.L. (2004b) A polyvalent DNA vaccine expressing an ESAT6-Ag85B fusion protein protects mice against a primary infection with *Mycobacterium tuberculosis* and boosts BCG-induced protective immunity. *Vaccine.* **23**, 780-788.
- Dietrich, J., Aagaard, C., Leah, R., Olsen, A.W., Stryhn, A., Doherty, T.M. and Andersen, P. (2005) Exchanging ESAT6 with TB10.4 in an Ag85B fusion molecule-based tuberculosis subunit vaccine: Efficient protection and ESAT6-based sensitive monitoring of vaccine efficacy. *J. Immunol.* **174**, 6332-6339.
- Dillon, D.C., Alderson, M.R., Day, C.H., Lewinson, D.M., Coler, R., Bement, T., Campos-Neto, A., Skeiky, Y.A.W., Orme, I.M., Roberts, A., Steen, S., Dalemans,

- W., Badaro, R. and Reed, S.G. (1999) Molecular characterization and human T-cell responses to a member of a novel *Mycobacterium tuberculosis* mtb39 gene family. *Infect. Immun.* **67**, 2941-2950.
- Dillon, D.C., Alderson, M.R., Day, C.H., Bement, T., Campos-Neto, A., Skeiky, Y.A.W., Vedvick, T., Badaro, R., Reed, S.G. and Houghton, R. (2000) Molecular and immunological characterization of *Mycobacterium tuberculosis* CFP-10, an immunodiagnostic antigen missing in *Mycobacterium bovis* BCG. *J. Clin. Microbiol.* **38**, 3285-3290.
- Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y.Z., Kilburn, A.E., Lee, W.H. and Elledge, S.J. (1993) The Retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.* **7**, 555-569.
- Dye, C., Scheele, S., Dolin, P., Pathania, V. and Raviglione, R.C. (1999) Global burden of tuberculosis - Estimated incidence, prevalence, and mortality by country. *JAMA.* **282**, 677-686.
- Eddine, A.N. and Kaufmann, S.H.E. (2005) Improved protection by recombinant BCG. *Microbes Infect.* **7**, 939-946.
- Elhay, M.J., Oettinger, T. and Andersen, P. (1998) Delayed-type hypersensitivity responses to ESAT-6 and MPT64 from *Mycobacterium tuberculosis* in the guinea pig. *Infect. Immun.* **66**, 3454-3456.
- Feilotter, H.E., Hannon, G.J., Ruddell, C.J. and Beach, D. (1994) Construction of an improved host strain for two-hybrid screening. *Nucleic Acids Res.* **22**, 1502-1503.
- Feizabadi, M.M., Robertson, I.D., Cousins, D.V. and Hampson, D.J. (1996) Genomic analysis of *Mycobacterium bovis* and other members of the *Mycobacterium tuberculosis* complex by isoenzyme analysis and pulsed-field gel electrophoresis. *J. Clin. Microbiol.* **34**, 1136-1142.
- Fenhalls, G., Wong, A., Bezuidenhout, J., van Helden, P., Bardin, P. and Lukey, P.T. (2000) *In situ* production of gamma interferon, interleukin-4, and tumor necrosis factor alpha mRNA in human lung tuberculous granulomas. *Infect. Immun.* **68**, 2827-2836.
- Fenhalls, G., Stevens-Muller, L., Warren, R., Carroll, N., Bezuidenhout, J., Van Helden, P. and Bardin, P. (2002a) Localisation of mycobacterial DNA and mRNA in human tuberculous granulomas. *J. Microbiol. Methods.* **51**, 197-208.
- Fenhalls, G., Stevens, L., Moses, L., Bezuidenhout, J., Betts, J.C., van Helden, P., Lukey, P.T. and Duncan, K. (2002b) *In situ* detection of *Mycobacterium tuberculosis* transcripts in human lung granulomas reveals differential gene expression in necrotic lesions. *Infect. Immun.* **70**, 6330-6338.
- Fenton, M.J. and Vermeulen, M.W. (1996) Immunopathology of tuberculosis: Roles of macrophages and monocytes. *Infect. Immun.* **64**, 683-690.
- Fields, S. and Song, O.K. (1989) A novel genetic system to detect protein-protein interactions. *Nature.* **340**, 245-246.
- Fine, P., Carneiro, I., Milstein, J. and Clements, C. (1999) Issues relating to the use of BCG in immunisation programmes. WHO/V&B/99.23, 1-42. Geneva
- Fleischmann, R.D., Alland, D., Eisen, J.A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J.F., Nelson, W.C., Umayam, L.A., Ermolaeva, M., Salzberg, S.L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A., Bishai, W., Jacobs, W.R., Venter, J.C. and Fraser, C.M. (2002) Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *J. Bacteriol.* **184**, 5479-5490.
- Flores, A.R., Parsons, L.M. and Pavelka, M.S. (2005) Genetic analysis of the β -lactamases of *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* and susceptibility to β -lactam antibiotics. *Microbiology.* **151**, 521-532.

- Flynn, J.L., Goldstein, M.M., Chan, J., Triebold, K.J., Pfeffer, K., Lowenstein, C.J., Schreiber, R., Mak, T.W. and Bloom, B.R. (1995) Tumor Necrosis Factor-Alpha is required in the protective immune response against *Mycobacterium tuberculosis* in mice. *Immunity*. **2**, 561-572.
- Flynn, J.L. and Chan, J. (2001) Tuberculosis: Latency and reactivation. *Infect. Immun.* **69**, 4195-4201.
- Fortune, S.M., Jaeger, A., Sarracino, D.A., Chase, M.R., Sasseti, C.M., Sherman, D.R., Bloom, B.R. and Rubin, E.J. (2005) Mutually dependent secretion of proteins required for mycobacterial virulence. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 10676-10681.
- Fratti, R.A., Backer, J.M., Gruenberg, J., Corvera, S. and Deretic, V. (2001) Role of phosphatidylinositol 3-kinase and Rab5 effectors in phagosomal biogenesis and mycobacterial phagosome maturation arrest. *J. Cell Biol.* **154**, 631-644.
- Fratti, R.A., Chua, J., Vergne, I. and Deretic, V. (2003) *Mycobacterium tuberculosis* glycosylated phosphatidylinositol causes phagosome maturation arrest. *Proc. Nat. Acad. Sci. U.S.A.* **100**, 5437-5442.
- Frota, C.C., Hunt, D.M., Buxton, R.S., Rickman, L., Hinds, J., Kremer, K., van Soolingen, D. and Colston, M.J. (2004) Genome structure in the vole bacillus, *Mycobacterium microti*, a member of the *Mycobacterium tuberculosis* complex with a low virulence for humans. *Microbiology*. **150**, 1519-1527.
- Frothingham, R., Hills, H.G. and Wilson, K.H. (1994) Extensive DNA sequence conservation throughout the *Mycobacterium tuberculosis* complex. *J. Clin. Microbiol.* **32**, 1639-1643.
- Frothingham, R., Strickland, P.L., Bretzel, G., Ramaswamy, S., Musser, J.M. and Williams, D.L. (1999) Phenotypic and genotypic characterization of *Mycobacterium africanum* isolates from West Africa. *J. Clin. Microbiol.* **37**, 1921-1926.
- Gao, L.Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J.N. and Brown, E.J. (2004) A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion. *Mol. Microbiol.* **53**, 1677-1693.
- Garcia, J., Puentes, A., Rodriguez, L., Ocampo, M., Curtidor, H., Vera, R., Lopez, R., Valbuena, J., Cortes, J., Vanegas, M., Barrero, C., Patarroyo, M.A., Urquiza, M. and Patarroyo, M.E. (2005) *Mycobacterium tuberculosis* Rv2536 protein implicated in specific binding to human cell lines. *Protein Sci.* **14**, 2236-2245.
- Garnier, T., Eiglmeier, K., Camus, J.C., Medina, N., Mansoor, H., Pryor, M., Duthoy, S., Grondin, S., Lacroix, C., Monsempe, C., Simon, S., Harris, B., Atkin, R., Doggett, J., Mayes, R., Keating, L., Wheeler, P.R., Parkhill, J., Barrell, B.G., Cole, S.T., Gordon, S.V. and Hewinson, R.G. (2003) The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 7877-7882.
- Gey van Pittius, N.C., Gamielien, J., Hide, W., Brown, G.D., Siezen, R.J. and Beyers, A.D. (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. *Genome Biol.* **2**, 0044.0041-0044.0018.
- Gietz, R.D. and Woods, R.A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Meth. Enzymol.* Vol. 350, pp. 87-96.
- Gil, D.P., Leon, L.G., Correa, L.I., Maya, J.R., Paris, S.C., Garcia, L.F. and Rojas, M. (2004) Differential induction of apoptosis and necrosis in monocytes from patients with tuberculosis and healthy control subjects. *J. Infect. Dis.* **189**, 2120-2128.
- Giniger, E., Varnum, S.M. and Ptashne, M. (1985) Specific DNA-binding of GAL4, a positive regulatory protein of yeast. *Cell.* **40**, 767-774.
- Gluzman, Y. (1981) SV40-transformed simian cells support the replication of early SV40 mutants. *Cell.* **23**, 175-182.

- Grode, L., Seiler, P., Baumann, S., Hess, J., Brinkmann, V., Eddine, A.N., Mann, P., Goosmann, C., Bandermann, S., Smith, D., Bancroft, G.J., Reyrat, J.M., van Soolingen, D., Raupach, B. and Kaufmann, S.H.E. (2005) Increased vaccine efficacy against tuberculosis of recombinant *Mycobacterium bovis* bacille Calmette-Guerin mutants that secrete listeriolysin. *J. Clin. Invest.* **115**, 2472-2479.
- Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S. and Sherman, D.R. (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol. Microbiol.* **51**, 359-370.
- Guleria, I., Teitelbaum, R., McAdam, R.A., Kalpana, G., Jacobs, W.R. and Bloom, B.R. (1996) Auxotrophic vaccines for tuberculosis. *Nat. Med.* **2**, 334-337.
- Harboe, M., Oettinger, T., Wiker, H.G., Rosenkrands, I. and Andersen, P. (1996) Evidence for occurrence of the ESAT-6 protein in *Mycobacterium tuberculosis* and virulent *Mycobacterium bovis* and for its absence in *Mycobacterium bovis* BCG. *Infect. Immun.* **64**, 16-22.
- Harboe, M., Wiker, H.G., Ulvund, G., Lund-Pedersen, B., Andersen, A.B., Hewinson, R.G. and Nagai, S. (1998) MPB70 and MPB83 as indicators of protein localization in mycobacterial cells. *Infect. Immun.* **66**, 289-296.
- Hart, P.D. and Sutherland, I. (1977) BCG and vole bacillus vaccines in prevention of tuberculosis in adolescence and early adult life: Final report to Medical Research Council. *Br. Med. J.* **2**, 293-295.
- Haslov, K., Andersen, A., Nagai, S., Gottschau, A., Sorensen, T. and Andersen, P. (1995) Guinea-pig cellular immune responses to proteins secreted by *Mycobacterium tuberculosis*. *Infect. Immun.* **63**, 804-810.
- Haugland, R., Spence, M., Johnson, I. and Basey, A. (2005) The Handbook - A Guide to Fluorescent Probes and Labelling Technologies. p11-17.
- Health Protection Agency. (2005) TB Update. *Health Protection Agency*, http://www.hpa.org.uk/infections/topics_az/tb/pdf/newsletter_2005.pdf. London.
- Hewinson, R.G. and Russell, W.P. (1993) Processing and secretion by *Escherichia coli* of a recombinant form of the immunogenic protein MPB70 of *Mycobacterium bovis*. *J. Gen. Microbiol.* **139**, 1253-1259.
- Hewinson, R.G., Michell, S.L., Russell, W.P., McAdam, R.A. and Jacobs, W.R. (1996) Molecular characterization of MPT83: A seroreactive antigen of *Mycobacterium tuberculosis* with homology to MPT70. *Scand. J. Immunol.* **43**, 490-499.
- Hogarth, P.J., Logan, K.E., Vordermeier, H.M., Singh, M., Hewinson, R.G. and Chambers, M.A. (2005) Protective immunity against *Mycobacterium bovis* induced by vaccination with Rv3109c: a member of the esat-6 gene family. *Vaccine.* **23**, 2557-2564.
- Hogarth, P.J., Logan, K. E., Ferraz, J. C., Hewinson, R. G. and Chambers M. A. (2006) Protective efficacy induced by *Mycobacterium bovis* bacille Calmette-Guerin can be augmented in an antigen independent manner by use of non-coding plasmid DNA. *Vaccine.* **24**, 95-101.
- Hope, I.A. and Struhl, K. (1986) Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell.* **46**, 885-894.
- Horwitz, M.A., Lee, B.W.E., Dillon, B.J. and Harth, G. (1995) Protective immunity against tuberculosis induced by vaccination with major extracellular proteins of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1530-1534.
- Hsu, T., Hingley-Wilson, S.M., Chen, B., Chen, M., Dai, A.Z., Morin, P.M., Marks, C.B., Padiyar, J., Goulding, C., Gingery, M., Eisenberg, D., Russell, R.G., Derrick, S.C., Collins, F.M., Morris, S.L., King, C.H. and Jacobs, W.R. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic

- function required for invasion of lung interstitial tissue. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12420-12425.
- Huygen, K. (2003) On the use of DNA vaccines for the prophylaxis of mycobacterial diseases. *Infect. Immun.* **71**, 1613-1621.
- Huygen, K. (2005) Plasmid DNA vaccination. *Microbes Infect.* **7**, 932-938.
- Imaeda, T. (1985) Deoxyribonucleic acid relatedness among selected strains of *Mycobacterium tuberculosis*, *Mycobacterium bovis*, *Mycobacterium bovis* BCG, *Mycobacterium microti*, and *Mycobacterium africanum*. *Int. J. Syst. Bacteriol.* **35**, 147-150.
- Inwald, J., Jahans, K., Hewinson, R.G. and Gordon, S.V. (2003) Inactivation of the *Mycobacterium bovis* homologue of the polymorphic RD1 gene Rv3879c (Mb3909c) does not affect virulence. *Tuberculosis.* **83**, 387-393.
- Jainchill, J., Aaronson, S. and Todaro, G. (1969) Murine sarcoma and leukemia viruses: Assay using clonal lines of contact-inhibited mouse cells. *J. Virol.* **4**, 549-553.
- James, P., Halladay, J. and Craig, E.A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics.* **144**, 1425-1436.
- Jarlier, V. and Nikaido, H. (1990) Permeability barrier to hydrophilic solutes in *Mycobacterium chelonae*. *J. Bacteriol.* **172**, 1418-1423.
- Johnston, M. (1987) A model fungal gene regulatory mechanism - the *GAL* genes of *Saccharomyces cerevisiae*. *Microbiol. Rev.* **51**, 458-476.
- Johnston, S.A., Salmeron, J.M. and Dincher, S.S. (1987) Interaction of positive and negative regulatory proteins in the galactose regulon of yeast. *Cell.* **50**, 143-146.
- Johnston, M. and Dover, J. (1987) Mutations that inactivate a yeast transcriptional regulatory protein cluster in an evolutionarily conserved DNA-binding domain. *Proc. Nat. Acad. Sci. U.S.A.* **84**, 2401-2405.
- Junge, S., Brenner, B., Lepple-Wienhues, A., Nilius, B., Lang, F., Linderkamp, O. and Gulbins, E. (1999) Intracellular mechanisms of L-selectin induced capping. *Cell. Signal.* **11**, 301-308.
- Kamath, A.T., Feng, C.G., Macdonald, M., Briscoe, H. and Britton, W.J. (1999) Differential protective efficacy of DNA vaccines expressing secreted proteins of *Mycobacterium tuberculosis*. *Infect. Immun.* **67**, 1702-1707.
- Kaufmann, S.H.E. (2005) Introduction. Rational vaccine development against tuberculosis: "Those who don't remember the past are condemned to repeat it". *Microbes Infect.* **7**, 897-898.
- Kaufmann, S.H.E. and McMichael, A.J. (2005) Annulling a dangerous liaison: vaccination strategies against AIDS and tuberculosis. *Nat. Med.* **11**, 578-578.
- Keane, J., Remold, H.G. and Kornfeld, H. (2000) Virulent *Mycobacterium tuberculosis* strains evade apoptosis of infected alveolar macrophages. *J. Immunol.* **164**, 2016-2020.
- Keegan, L., Gill, G. and Ptashne, M. (1986) Separation of DNA-binding from the transcription-activating function of a eukaryotic regulatory protein. *Science.* **231**, 699-704.
- Khan, A.A., Steiner, J.P., Klein, M.G., Schneider, M.F. and Snyder, S.H. (1992) IP3 receptor - Localization to plasma membrane of T-cells and cocapping with the T-cell receptor. *Science.* **257**, 815-818.
- Kishore, G.M. and Shah, D.M. (1988) Amino acid biosynthesis inhibitors as herbicides. *Annu. Rev. Biochem.* **57**, 627-663.
- Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55.
- Kwiatkowska, K. and Sobota, A. (1999a) Signaling pathways in phagocytosis. *Bioessays.* **21**, 422-431.

- Kwiatkowska, K. and Sobota, A. (1999b) Tyrosine phosphorylation/dephosphorylation controls capping of Fcγ receptor II in U937 cells. *Cell Motil. Cytoskeleton.* **42**, 298-314.
- Ladokhin, A.S. (2000) Fluorescence Spectroscopy in Peptide and Protein Analysis. In Meyers, R.A. (ed.), *Encyclopedia of Analytical Chemistry*. John Wiley & Sons Ltd, Chichester, pp. 5672-5779.
- Lakowicz, J.R. (1983) Protein Fluorescence. In *Principles of Fluorescence Spectroscopy*. Plenum Press, New York, pp. 341-382.
- Langermans, J.A.M., Doherty, T.M., Vervenne, R.A.W., van der Laan, T., Lyashchenko, K., Greenwald, R., Agger, E.M., Aagaard, C., Weiler, H., van Soolingen, D., Dalemans, W., Thomas, A.W. and Andersen, P. (2005) Protection of macaques against *Mycobacterium tuberculosis* infection by a subunit vaccine based on a fusion protein of antigen 85B and ESAT-6. *Vaccine.* **23**, 2740-2750.
- Laughon, A. and Gesteland, R.F. (1984) Primary structure of the *Saccharomyces cerevisiae* GAL4 gene. *Mol. Cell. Biol.* **4**, 260-267.
- Le Douarin, B., Heery, D. M., Gaudon, C., vom Baur, E. and Losson R. (2001) Yeast two-hybrid screening for proteins that interact with nuclear hormone receptors. In *Methods in Molecular Biology*. Vol. 176, pp. 227-248.
- Lewis, K.N., Liao, R.L., Guinn, K.M., Hickey, M.J., Smith, S., Behr, M.A. and Sherman, D.R. (2003) Deletion of RD1 from *Mycobacterium tuberculosis* mimics bacille Calmette-Guerin attenuation. *J. Infect. Dis.* **187**, 117-123.
- Li, Z., Song, D., Zhang, H., He, W., Fan, X., Zhang, Y., Huang, J., Wang, X., Liu, Q. and Xiong, S. (2006) Improved humoral immunity against tuberculosis ESAT-6 antigen by chimeric DNA prime and protein boost strategy. *DNA Cell Biol.* **25**, 25-30.
- Liao, F., Shin, H.S. and Rhee, S.G. (1992) Tyrosine phosphorylation of phospholipase C-γ1 induced by cross-linking of the high-affinity or low-affinity Fc receptor for IgG in U937 cells. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3659-3663.
- Lightbody, K.L., Renshaw, P.S., Collins, M.L., Wright, R.L., Hunt, D.M., Gordon, S.V., Hewinson, R.G., Buxton, R.S., Williamson, R.A. and Carr, M.D. (2004) Characterisation of complex formation between members of the *Mycobacterium tuberculosis* complex CFP-10/ESAT-6 protein family: towards an understanding of the rules governing complex formation and thereby functional flexibility. *FEMS Microbiol. Lett.* **238**, 255-262.
- Ma, J. and Ptashne, M. (1987a) The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80. *Cell.* **50**, 137-142.
- Ma, J. and Ptashne, M. (1987b) Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell.* **48**, 847-853.
- Ma, J. and Ptashne, M. (1988) Converting a eukaryotic transcriptional inhibitor into an activator. *Cell.* **55**, 443-446.
- MacGurn, J.A., Raghavan, S., Stanley, S.A. and Cox, J.S. (2005) A non-RD1 gene cluster is required for Snm secretion in *Mycobacterium tuberculosis*. *Mol. Microbiol.* **57**, 1653-1663.
- Mahairas, G.G., Sabo, P.J., Hickey, M.J., Singh, D.C. and Stover, C.K. (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J. Bacteriol.* **178**, 1274-1282.
- Majlessi, L., Brodin, P., Brosch, R., Rojas, M.J., Khun, H., Huerre, M., Cole, S.T. and Leclerc, C. (2005) Influence of ESAT-6 secretion system 1 (RD1) of *Mycobacterium tuberculosis* on the interaction between mycobacteria and the host immune system. *J. Immunol.* **174**, 3570-3579.
- Malik, Z.A., Denning, G.M. and Kusner, D.J. (2000) Inhibition of Ca²⁺ signaling by *Mycobacterium tuberculosis* is associated with reduced phagosome-lysosome

- fusion and increased survival within human macrophages. *J. Exp. Med.* **191**, 287-302.
- Malik, Z.A., Thompson, C.R., Hashimi, S., Porter, B., Iyer, S.S. and Kusner, D.J. (2003) *Mycobacterium tuberculosis* blocks Ca²⁺ signaling and phagosome maturation in human macrophages via specific inhibition of sphingosine kinase. *J. Immunol.* **170**, 2811-2815.
- Marei, A., Ghaemmaghami, A., Renshaw, P., Wiselka, M., Barer, M., Carr, M. and Ziegler-Heitbrock, L. (2005) Superior T cell activation by ESAT-6 as compared with the ESAT-6-CFP-10 complex. *Int. Immunol.* **17**, 1439-1446.
- Mattow, J., Schaible, U.E., Schmidt, F., Hagens, K., Siejak, R., Brestrich, G., Haeselbarth, G., Muller, E.C., Jungblut, P.R. and Kaufmann, S.H.E. (2003) Comparative proteome analysis of culture supernatant proteins from virulent *Mycobacterium tuberculosis* H37Rv and attenuated *M. bovis* BCG Copenhagen. *Electrophoresis.* **24**, 3405-3420.
- McAdam, R.A., Weisbrod, T.R., Martin, J., Scuderi, J.D., Brown, A.M., Cirillo, J.D., Bloom, B.R. and Jacobs, W.R. (1995) *In Vivo* growth characteristics of leucine and methionine auxotrophic mutants of *Mycobacterium bovis* BCG generated by transposon mutagenesis. *Infect. Immun.* **63**, 1004-1012.
- Mdluli, K., Slayden, R.A., Zhu, Y.Q., Ramaswamy, S., Pan, X., Mead, D., Crane, D.D., Musser, J.M. and Barry, C.E. (1998) Inhibition of a *Mycobacterium tuberculosis* beta-ketoacyl ACP synthase by isoniazid. *Science.* **280**, 1607-1610.
- Michell, S.L., Whelan, A.O., Wheeler, P.R., Partico, M., Easton, R.L., Etienne, A.T., Haslam, S.M., Dell, A., Morris, H.R., Reason, A.J., Herrmann, J.L., Young, D.B. and Hewinson, R.G. (2003) The MPB83 antigen from *Mycobacterium bovis* contains O-linked mannose and (1→3)-mannobiose moieties. *J. Biol. Chem.* **278**, 16423-16432.
- Miura, K., Nagai, S., Kinomoto, M., Haga, S. and Tokunaga, T. (1983) Comparative studies with various substrains of *Mycobacterium bovis* BCG on the production of an antigenic protein, MPB70. *Infect. Immun.* **39**, 540-545.
- Mostowy, S., Tsolaki, A.G., Small, P.M. and Behr, M.A. (2003) The *in vitro* evolution of BCG vaccines. *Vaccine.* **21**, 4270-4274.
- Murray, P.J., Aldovini, A. and Young, R.A. (1996) Manipulation and potentiation of antimycobacterial immunity using recombinant bacille Calmette-Guerin strains that secrete cytokines. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 934-939.
- Musser, J.M., Kapur, V., Williams, D.L., Kreiswirth, B.N., van Soolingen, D. and van Embden, J.D.A. (1996) Characterization of the catalase-peroxidase gene (*katG*) and *inhA* locus in isoniazid-resistant and -susceptible strains of *Mycobacterium tuberculosis* by automated DNA sequencing: Restricted array of mutations associated with drug resistance. *J. Infect. Dis.* **173**, 196-202.
- Musser, J.M., Amin, A. and Ramaswamy, S. (2000) Negligible genetic diversity of *Mycobacterium tuberculosis* host immune system protein targets: Evidence of limited selective pressure. *Genetics.* **155**, 7-16.
- Mustafa, A.S., Amoudy, H.A., Wiker, H.G., Abal, A.T., Ravn, P., Oftung, F. and Andersen, P. (1998) Comparison of antigen-specific T-cell responses of tuberculosis patients using complex or single antigens of *Mycobacterium tuberculosis*. *Scand. J. Immunol.* **48**, 535-543.
- Mustafa, A.S., Cockle, P.J., Shaban, F., Hewinson, R.G. and Vordermeier, H.M. (2002) Immunogenicity of *Mycobacterium tuberculosis* RD1 region gene products in infected cattle. *Clin. Exp. Immunol.* **130**, 37-42.
- Niemann, S., Richter, E., Dalugge-Tamm, H., Schlesinger, H., Graupner, D., Konigstein, B., Gurath, G., Greinert, U. and Rusch-Gerdes, S. (2000) Two cases of

- Mycobacterium microti* derived tuberculosis in HIV-negative immunocompetent patients. *Emerg. Infect. Dis.*, **6**, 539-542.
- Okkels, L.M., Brock, I., Follmann, F., Agger, E.M., Arend, S.M., Ottenhoff, T.H.M., Oftung, F., Rosenkrands, I. and Andersen, P. (2003) PPE protein (Rv3873) from DNA segment RD1 of *Mycobacterium tuberculosis*: Strong recognition of both specific T-cell epitopes and epitopes conserved within the PPE family. *Infect. Immun.* **71**, 6116-6123.
- Okkels, L.M. and Andersen, P. (2004) Protein-protein interactions of proteins from the ESAT-6 family of *Mycobacterium tuberculosis*. *J. Bacteriol.* **186**, 2487-2491.
- Okkels, L.M., Muller, E.C., Schmid, M., Rosenkrands, I., Kaufmann, S.H.E., Andersen, P. and Jungblut, P.R. (2004) CFP10 discriminates between nonacetylated and acetylated ESAT-6 of *Mycobacterium tuberculosis* by differential interaction. *Proteomics.* **4**, 2954-2960.
- Olsen, A.W., van Pinxteren, L.A.H., Okkels, L.M., Rasmussen, P.B. and Andersen, P. (2001) Protection of mice with a tuberculosis subunit vaccine based on a fusion protein of antigen 85B and ESAT-6. *Infect. Immun.* **69**, 2773-2778.
- Olsen, A.W., Williams, A., Okkels, L.M., Hatch, G. and Andersen, P. (2004) Protective effect of a tuberculosis subunit vaccine based on a fusion of antigen 85B and ESAT-6 in the aerosol guinea pig model. *Infect. Immun.* **72**, 6148-6150.
- Orme, I.M. (1988) Characteristics and specificity of acquired immunological memory to *Mycobacterium tuberculosis* infection. *J. Immunol.* **140**, 3589-3593.
- Pai, M., Riley, L.W. and Colford, J.M. (2004) Interferon- γ assays in the immunodiagnosis of tuberculosis: a systematic review. *Lancet Infect. Dis.* **4**, 761-776.
- Palendira, U., Spratt, J.M., Britton, W.J. and Triccas, J.A. (2005) Expanding the antigenic repertoire of BCG improves protective efficacy against aerosol *Mycobacterium tuberculosis* infection. *Vaccine.* **23**, 1680-1685.
- Pallen, M.J. (2002) The ESAT-6/WXG100 superfamily - and a new Gram positive secretion system? *Trends Microbiol.* **10**, 209-212.
- Panchuk-Voloshina, N., Haugland, R.P., Bishop-Stewart, J., Bhalgat, M.K., Millard, P.J., Mao, F., Leung, W.Y. and Haugland, R.P. (1999) Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J. Histochem. Cytochem.* **47**, 1179-1188.
- Pasca, M.R., Gugliera, P., Arcesi, F., Bellinzoni, M., De Rossi, E. and Riccardi, G. (2004) Rv2686c-Rv2687c-Rv2688c, an ABC fluoroquinolone efflux pump in *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **48**, 3175-3178.
- Philipp, W.J., Nair, S., Guglielmi, G., Lagranderie, M., Gicquel, B. and Cole, S.T. (1996) Physical mapping of *Mycobacterium bovis* BCG Pasteur reveals differences from the genome map of *Mycobacterium tuberculosis* H37Rv and from *M. bovis*. *Microbiology.* **142**, 3135-3145.
- Placido, R., Mancino, G., Amendola, A., Mariani, F., Vendetti, S., Piacentini, M., Sanduzzi, A., Bocchino, M.L., Zembala, M. and Colizzi, V. (1997) Apoptosis of human monocytes/macrophages in *Mycobacterium tuberculosis* infection. *J. Path.* **181**, 31-38.
- Pym, A.S., Brodin, P., Brosch, R., Huerre, M. and Cole, S.T. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol. Microbiol.* **46**, 709-717.
- Pym, A.S., Brodin, P., Majlessi, L., Brosch, R., Demangel, C., Williams, A., Griffiths, K.E., Marchal, G., Leclerc, C. and Cole, S.T. (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat. Med.* **9**, 533-539.
- Ramaswamy, S. and Musser, J.M. (1998) Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*. *Tuber. Lung Dis.* **79**, 3-29.
- Raviglione, M.C. (2003) The TB epidemic from 1992 to 2002. *Tuberculosis.* **83**, 4-14.

- Rattan, A., Kalia, A. and Ahmad, N. (1998) Multidrug resistant *Mycobacterium tuberculosis*: Molecular perspectives. *Emerg. Infect. Dis.* **4**, 195-209.
- Ravn, P., Demissie, A., Eguale, T., Wondwosson, H., Lein, D., Amoudy, H.A., Mustafa, A.S., Jensen, A.K., Holm, A., Rosenkrands, I., Oftung, F., Olobo, J., von Reyn, F. and Andersen, P. (1999) Human T cell responses to the ESAT-6 antigen from *Mycobacterium tuberculosis*. *J. Infect. Dis.* **179**, 637-645.
- Renshaw, P.S., Panagiotidou, P., Whelan, A., Gordon, S.V., Hewinson, R.G., Williamson, R.A. and Carr, M.D. (2002) Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1: 1 complex and characterization of the structural properties of ESAT-6, CFP-10, and the ESAT-6•CFP-10 complex - Implications for pathogenesis and virulence. *J. Biol. Chem.* **277**, 21598-21603.
- Renshaw, P.S., Lightbody, K.L., Veverka, V., Muskett, F.W., Kelly, G., Frenkiel, T.A., Gordon, S.V., Hewinson, R.G., Burke, B., Norman, J., Williamson, R.A. and Carr, M.D. (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J.* **24**, 2491-2498.
- Rickman, L., Scott, C., Hunt, D.M., Hutchinson, T., Menendez, M.C., Whalan, R., Hinds, J., Colston, M.J., Green, J. and Buxton, R.S. (2005) A member of the cAMP receptor protein family of transcription regulators in *Mycobacterium tuberculosis* is required for virulence in mice and controls transcription of the *rpfA* gene coding for a resuscitation promoting factor. *Mol. Microbiol.* **56**, 1274-1286.
- Rojas, M., Olivier, M., Gros, P., Barrera, L.F. and Garcia, L.F. (1999) TNF α and IL-10 modulate the induction of apoptosis by virulent *Mycobacterium tuberculosis* in murine macrophages. *J. Immunol.* **162**, 6122-6131.
- Romain, F., Laqueyrie, A., Militzer, P., Pescher, P., Chavarot, P., Lagranderie, M., Auregan, G., Gheorghiu, M. and Marchal, G. (1993) Identification of a *Mycobacterium bovis* BCG 45/47-kilodalton antigen complex, an immunodominant target for antibody response after immunization with living bacteria. *Infect. Immun.* **61**, 742-750.
- Rosenkrands, I., Weldingh, K., Ravn, P., Brandt, L., Hojrup, P., Rasmussen, P.B., Coates, A.R., Singh, M., Mascagni, P. and Andersen, P. (1999) Differential T-cell recognition of native and recombinant *Mycobacterium tuberculosis* GroES. *Infect. Immun.* **67**, 5552-5558.
- Rosenkrands, I., Weldingh, K., Jacobsen, S., Hansen, C.V., Florio, W., Gianetri, I. and Andersen, P. (2000) Mapping and identification of *Mycobacterium tuberculosis* proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection. *Electrophoresis.* **21**, 935-948.
- Rupp, S. (2002) *LacZ* assays in yeast. *Meth. Enzymol.* Vol. 350, pp. 112-131.
- Sambandamurthy, V.K. and Jacobs, W.R. (2005) Live attenuated mutants of *Mycobacterium tuberculosis* as candidate vaccines against tuberculosis. *Microbes Infect.* **7**, 955-961.
- Sampson, S.L., Dascher, C.C., Sambandamurthy, V.K., Russell, R.G., Jacobs, W.R., Bloom, B.R. and Hondalus, M.K. (2004) Protection elicited by a double leucine and pantothenate auxotroph of *Mycobacterium tuberculosis* in guinea pigs. *Infect. Immun.* **72**, 3031-3037.
- Sasseti, C.M. and Rubin, E.J. (2003) Genetic requirements for mycobacterial survival during infection. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 12989-12994.
- Schnappinger, D., Ehrt, S., Voskuil, M.I., Liu, Y., Mangan, J.A., Monahan, I.M., Dolganov, G., Efron, B., Butcher, P.D., Nathan, C. and Schoolnik, G.K. (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: Insights into the phagosomal environment. *J. Exp. Med.* **198**, 693-704.

- Serebriiskii, I.G. and Golemis, E.A. (2000) Uses of lacZ to study gene function: Evaluation of β -galactosidase assays employed in the yeast two-hybrid system. *Anal. Biochem.* **285**, 1-15.
- Sherman, D.R., Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L. and Smith, S. (2004) *Mycobacterium tuberculosis* H37Rv: Δ RD1 is more virulent than *M. bovis* bacille Calmette-Guerin in long-term murine infection. *J. Infect. Dis.* **190**, 123-126.
- Silva, P.E.A., Bigi, F., Santangelo, M.D., Romano, M.I., Martin, C., Cataldi, A. and Ainsa, J.A. (2001) Characterization of P55, a multidrug efflux pump in *Mycobacterium bovis* and *Mycobacterium tuberculosis*. *Antimicrob. Agents Chemother.* **45**, 800-804.
- Singh, G., Singh, B., Vladimir, T. and Sharma, P. (2005) *Mycobacterium tuberculosis* 6 kDa early secreted antigenic target stimulates activation of J774 macrophages. *Immunol. Lett.* **98**, 180-188.
- Skeiky, Y.A.W., Ovendale, P.J., Jen, S., Alderson, M.R., Dillon, D.C., Smith, S., Wilson, C.B., Orme, I.M., Reed, S.G. and Campos-Neto, A. (2000) T cell expression cloning of a *Mycobacterium tuberculosis* gene encoding a protective antigen associated with the early control infection. *J. Immunol.* **165**, 7140-7149.
- Skinner, M.A., Buddle, B.M., Wedlock, D.N., Keen, D., de Lisle, G.W., Tascon, R.E., Ferraz, J.C., Lowrie, D.B., Cockle, P.J., Vordermeier, H.M. and Hewinson, R.G. (2003) A DNA prime-*Mycobacterium bovis* BCG boost vaccination strategy for cattle induces protection against bovine tuberculosis. *Infect. Immun.* **71**, 4901-4907.
- Skjot, R.L.V., Oettinger, T., Rosenkrands, I., Ravn, P., Brock, I., Jacobsen, S. and Andersen, P. (2000) Comparative evaluation of low-molecular-mass proteins from *Mycobacterium tuberculosis* identifies members of the ESAT-6 family as immunodominant T-cell antigens. *Infect. Immun.* **68**, 214-220.
- Skjot, R.L.V., Brock, I., Arend, S.M., Munk, M.E., Theisen, M., Ottenhoff, T.H.M. and Andersen, P. (2002) Epitope mapping of the immunodominant antigen TB10.4 and the two homologous proteins TB10.3 and TB12.9, which constitute a subfamily of the esat-6 gene family. *Infect. Immun.* **70**, 5446-5453.
- Slayden, R.A. and Barry, C.E. (2000) The genetics and biochemistry of isoniazid resistance in *Mycobacterium tuberculosis*. *Microb. Infect.* **2**, 659-669.
- Smith, D.A., Parish, T., Stoker, N.G. and Bancroft, G.J. (2001) Characterization of auxotrophic mutants of *Mycobacterium tuberculosis* and their potential as vaccine candidates. *Infect. Immun.* **69**, 1142-1150.
- Smith, N.H., Kremer, K., Inwald, J., Dale, J., Driscoll, J.R., Gordon, S.V., van Soolingen, D., Hewinson, R.G. and Smith, J.M. (2005) Ecotypes of the *Mycobacterium tuberculosis* complex. *J. Theor. Biol.* In Press.
- Snapper, C.M., Yamada, H., Mond, J.J. and June, C.H. (1991) Cross-linkage of Ly-6a/E Induces Ca^{2+} translocation in the absence of phosphatidylinositol turnover and mediates proliferation of normal murine B lymphocytes. *J. Immunol.* **147**, 1171-1179.
- Sorensen, A.L., Nagai, S., Houen, G., Andersen, P. and Andersen, A.B. (1995) Purification and characterization of a low molecular mass T-cell antigen secreted by *Mycobacterium tuberculosis*. *Infect. Immun.* **63**, 1710-1717.
- Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S. and Musser, J.M. (1997) Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Nat. Acad. Sci. U.S.A.* **94**, 9869-9874.
- Stanley, S.A., Raghavan, S., Hwang, W.W. and Cox, J.S. (2003) Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialized secretion system. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 13001-13006.

- Stenger, S., Hanson, D.A., Teitelbaum, R., Dewan, P., Niazi, K.R., Froelich, C.J., Ganz, T., Thoma-Uszynski, S., Melian, A., Bogdan, C., Porcelli, S.A., Bloom, B.R., Krensky, A.M. and Modlin, R.L. (1998) An antimicrobial activity of cytolytic T cells mediated by granulysin. *Science*. **282**, 121-125.
- Stenger, S. and Modlin, R.L. (1999) T-cell mediated immunity to *Mycobacterium tuberculosis*. *Curr. Opin. Microbiol.* **2**, 89-93.
- Sturgill-Koszycki, S., Schlesinger, P.H., Chakraborty, P., Haddix, P.L., Collins, H.L., Fok, A.K., Allen, R.D., Gluck, S.L., Heuser, J. and Russell, D.G. (1994) Lack of acidification in *Mycobacterium* phagosomes produced by exclusion of the vesicular proton ATPase. *Science*. **263**, 678-681.
- Sundstrom, C. and Nilsson, K. (1976) Establishment and characterization of a human histiocytic lymphoma cell line (U-937). *Int. J. Cancer*. **17**, 565-577.
- Takiff, H.E., Cimino, M., Musso, M.C., Weisbrod, T., Martinez, R., Delgado, M.B., Salazar, L., Bloom, B.R. and Jacobs, W.R. (1996) Efflux pump of the proton antiporter family confers low-level fluoroquinolone resistance in *Mycobacterium smegmatis*. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 362-366.
- Health Protection Agency Centre for Infections, TB Section. (2005) Annual report on tuberculosis cases reported in England, Wales and Northern Ireland in 2003. http://www.hpa.org.uk/infections/topics_az/tb/pdf/2003_Annual_Report.pdf. London.
- Tekaia, F., Gordon, S., Garnier, T., Brosch, R., Barrell, B. and Cole, S. (1999) Analysis of the proteome of *Mycobacterium tuberculosis* in silico. *Tuber. Lung Dis.* **79**, 329-342.
- Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M.J., Matter, L., Schopfer, K. and Bodmer, T. (1993) Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. *Lancet*. **341**, 647-650.
- The Report of the Chief Veterinary Officer. (2005) Animal Health 2004. *Department for Environment, Food and Rural Affairs; Scottish Executive Environment and Rural Affairs Department; Welsh Assembly Government*, <http://www.defra.gov.uk/corporate/publications/pubcat/cvo/2004/index.htm>.
- Tirode, F., Malaguti, C., Romero, F., Attar, R., Camonis, J. and Egly, J.M. (1997) A conditionally expressed third partner stabilizes or prevents the formation of a transcriptional activator in a three-hybrid system. *J. Biol. Chem.* **272**, 22995-22999.
- Tollefsen, S., Vordermeier, M., Olsen, I., Storset, A.K., Reitan, L.J., Clifford, D., Lowrie, D.B., Wiker, H.G., Huygen, K., Hewinson, G., Mathiesen, I. and Tjelle, T.E. (2003) DNA injection in combination with electroporation: A novel method for vaccination of farmed ruminants. *Scand. J. Immunol.* **57**, 229-238.
- Trajkovic, V., Singh, G., Singh, B., Singh, S. and Sharma, P. (2002) Effect of *Mycobacterium tuberculosis* specific 10-kilodalton antigen on macrophage release of tumor necrosis factor alpha and nitric oxide. *Infect. Immun.* **70**, 6558-6566.
- Ulrichs, T., Kosmiadi, G.A., Trusov, V., Jorg, S., Pradl, L., Titukhina, M., Mishenko, V., Gushina, N. and Kaufmann, S.H.E. (2004) Human tuberculous granulomas induce peripheral lymphoid follicle-like structures to orchestrate local host defence in the lung. *J. Path.* **204**, 217-228.
- Van Criekinge, W. and R, B. (1999) Yeast two-hybrid: State of the art. *Biol. Proced. Online*. **2**, 1-38.
- Van Pinxteren, L.A.H., Ravn, P., Agger, E.M., Pollock, J. and Andersen, P. (2000) Diagnosis of tuberculosis based on the two specific antigens ESAT-6 and CFP10. *Clin. Diagn. Lab. Immunol.* **7**, 155-160.
- Van Soolingen, D., van der Zanden, A.G.M., de Haas, P.E.W., Noordhoek, G.T., Kiers, A., Foudraïne, N.A., Portaels, F., Kolk, A.H.J., Kremer, K. and van Embden, J.D.A.

- (1998) Diagnosis of *Mycobacterium microti* infections among humans by using novel genetic markers. *J. Clin. Microbiol.* **36**, 1840-1845.
- Van Soolingen, D., Hoogenboezem, T., deHaas, P.E.W., Hermans, P.W.M., Koedam, M.A., Teppema, K.S., Brennan, P.J., Besra, G.S., Portaels, F., Top, J., Schouls, L.M. and van Embden, J.D.A. (1997) A novel pathogenic taxon of the *Mycobacterium tuberculosis* complex, Canetti: Characterization of an exceptional isolate from Africa. *Int. J. Syst. Bacteriol.* **47**, 1236-1245.
- Vergne, I., Fratti, R.A., Hill, P.J., Chua, J., Belisle, J. and Deretic, V. (2004) *Mycobacterium tuberculosis* phagosome maturation arrest: mycobacterial phosphatidylinositol analog phosphatidylinositol mannoside stimulates early endosomal fusion. *Mol. Cell Biol.* **15**, 751-760.
- Via, L.E., Deretic, D., Ulmer, R.J., Hibler, N.S., Huber, L.A. and Deretic, V. (1997) Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by Rab5 and Rab7. *J. Biol. Chem.* **272**, 13326-13331.
- Volkman, H.E., Clay, H., Beery, D., Chang, J.C.W., Sherman, D.R. and Ramakrishnan, L. (2004) Tuberculous granuloma formation is enhanced by a mycobacterium virulence determinant. *PLoS Biology.* **2**, 1946-1956.
- Vordermeier, H.M., Chambers, M.A., Cockle, P.J., Whelan, A.O., Simmons, J. and Hewinson, R.G. (2002) Correlation of ESAT-6-specific gamma interferon production with pathology in cattle following *Mycobacterium bovis* BCG vaccination against experimental bovine tuberculosis. *Infect. Immun.* **70**, 3026-3032.
- Voskuil, M.I., Schnappinger, D., Rutherford, R., Liu, Y. and Schoolnik, G.K. (2004) Regulation of the *Mycobacterium tuberculosis* PE/PPE genes. *Tuberculosis.* **84**, 256-262.
- Vossebeld, P.J.M., Kessler, J., Vondemorne, A., Roos, D. and Verhoeven, A.J. (1995) Heterotypic FcγR clusters evoke a synergistic Ca²⁺ response in human neutrophils. *J. Biol. Chem.* **270**, 10671-10679.
- Wards, B.J., de Lisle, G.W. and Collins, D.M. (2000) An esat6 knockout mutant of *Mycobacterium bovis* produced by homologous recombination will contribute to the development of a live tuberculosis vaccine. *Tuber. Lung Dis.* **80**, 185-189.
- Wedlock, D.N., Skinner, M.A., Parlane, N.A., Vordermeier, H.M., Hewinson, R.G., de Lisle, G.W. and Buddle, B.M. (2003) Vaccination with DNA vaccines encoding MPB70 or MPB83 or a MPB70 DNA prime-protein boost does not protect cattle against bovine tuberculosis. *Tuberculosis.* **83**, 339-349.
- World Health Organisation. (2004a) TB/HIV A Clinical Manual. *WHO Report 2004*. WHO/HTM/TB/2004.329. Geneva.
- World Health Organisation. (2004b) Anti-tuberculosis drug resistance in the world. *WHO Report 2004*. WHO/HTM/TB/2004.343. Geneva.
- World Health Organisation. (2005) Global tuberculosis control - surveillance, planning, financing. *WHO Report 2005*. WHO/HTM/TB/2005. Geneva.
- Wiker, H.G. and Harboe, M. (1992) The antigen 85 complex: A major secretion product of *Mycobacterium tuberculosis*. *Microbiol. Rev.* **56**, 648-661.
- Wiker, H.G., Nagai, S., Hewinson, R.G., Russell, W.P. and Harboe, M. (1996) Heterogenous expression of the related MBP70 and MPB83 proteins distinguish various substrains of *Mycobacterium bovis* BCG and *Mycobacterium tuberculosis* H37Rv. *Scand. J. Immunol.* **43**, 374-380.
- Zhang, J. and Lautar, S. (1996) A yeast three-hybrid method to clone ternary protein complex components. *Anal. Biochem.* **242**, 68-72.
- Zhang, J., Jiang, R., Takayama, H. and Tanaka, Y. (2005) Survival of virulent *Mycobacterium tuberculosis* involves preventing apoptosis induced by Bcl-2

upregulation and release resulting from necrosis in J774 macrophages. *Microbiol. Immunol.* **49**, 845-852.

Zhang, Y., Heym, B., Allen, B., Young, D. and Cole, S. (1992) The catalase peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature.* **358**, 591-593.

Ziegler-Heitbrock, H.W.L., Thiel, E., Futterer, A., Herzog, V., Wirtz, A. and Riethmuller, G. (1988) Establishment of a human cell-line (Mono Mac 6) with characteristics of mature monocytes. *Int. J. Cancer.* **41**, 456-461.

Published Work

Lightbody, K.L., Renshaw, P.S., Collins, M.L., Wright, R.L., Hunt, D.M., Gordon, S.V., Hewinson, R.G., Buxton, R.S., Williamson, R.A. and Carr, M.D. (2004) Characterisation of complex formation between members of the *Mycobacterium tuberculosis* complex CFP-10/ESAT-6 protein family: towards an understanding of the rules governing complex formation and thereby functional flexibility. *FEMS Microbiol. Lett.* 238, 255-262.

Renshaw, P.S., Lightbody, K.L., Veverka, V., Muskett, F.W., Kelly, G., Frenkiel, T.A., Gordon, S.V., Hewinson, R.G., Burke, B., Norman, J., Williamson, R.A. and Carr, M.D. (2005) Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6. *EMBO J.* 24, 2491-2498.

Characterisation of complex formation between members of the *Mycobacterium tuberculosis* complex CFP-10/ESAT-6 protein family: towards an understanding of the rules governing complex formation and thereby functional flexibility

Kirsty L. Lightbody^a, Philip S. Renshaw^a, Michelle L. Collins^b, Rebecca L. Wright^a,
Debbie M. Hunt^c, Stephen V. Gordon^d, R. Glyn Hewinson^d, Roger S. Buxton^c,
Richard A. Williamson^b, Mark D. Carr^{a,*}

^a Department of Biochemistry, University of Leicester, Adrian Building, University Road, Leicester LE1 7RH, UK

^b Department of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, UK

^c Division of Mycobacterial Research, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

^d TB Research Group, Veterinary Laboratories Agency, Weybridge, New Haw, Addlestone, Surrey KT15 3NB, UK

Received 30 April 2004; received in revised form 5 July 2004; accepted 19 July 2004

First published online 29 July 2004

Abstract

We have previously shown that the secreted *M. tuberculosis* complex proteins CFP-10 and ESAT-6 form a tight, 1:1 complex, which may represent their functional form. In the work reported here a combination of yeast two-hybrid and biochemical analysis has been used to characterise complex formation between two other pairs of CFP-10/ESAT-6 family proteins (Rv0287/Rv0288 and Rv3019c/Rv3020c) and to determine whether complexes can be formed between non-genome paired members of the family. The results clearly demonstrate that Rv0287/Rv0288 and Rv3019c/3020c form tight complexes, as initially observed for CFP-10/ESAT-6. The closely related Rv0287/Rv0288 and Rv3019c/Rv3020c proteins are also able to form non-genome paired complexes (Rv0287/Rv3019c and Rv0288/Rv3020c), but are not capable of binding to the more distantly related CFP-10/ESAT-6 proteins. © 2004 Federation of European Microbiological Societies. Published by Elsevier B.V. All rights reserved.

Keywords: CFP-10; ESAT-6; Tuberculosis; Protein–protein interaction; ESAT-6 family members

1. Introduction

An estimated 2 billion people are currently infected by *Mycobacterium tuberculosis* bacilli, resulting in the deaths of approximately 2 million people annually. The genome of *Mycobacterium tuberculosis* H37Rv contains the genes for just over 4000 proteins, only about

half of which have been ascribed any specific function [1]. Consequently, we still have very little information about which proteins are essential for tuberculosis pathogenesis and even less knowledge of their structures, functions and mechanisms of action.

Comparative genomic studies identified a total of 14 regions of difference between the *M. bovis* BCG-Pasteur vaccine strain and virulent *M. tuberculosis* [2]. One region, designated RD1, is deleted from all BCG sub-strains but is present in all clinical isolates and virulent strains of *M. bovis* and *M. tuberculosis* [2]. The RD1

* Corresponding author. Tel.: +44 116 252 3054; fax: +44 116 223 1503.

E-mail address: mdc12@le.ac.uk (M.D. Carr).

region contains the genes for 9 proteins, with open reading frame designations Rv3871 to Rv3879c, which are clearly implicated in tuberculosis pathogenesis [2,3]. This region includes the genes for two, sequence related, relatively small proteins, which are known as CFP-10 (encoded by Rv3874 or *esxB*) and ESAT-6 (encoded by Rv3875 or *esxA*). The expression of these two genes is co-ordinately regulated [4] and the proteins appear to be secreted by a membrane protein complex formed from the products of several flanking genes [5–7]. The importance of CFP-10 and ESAT-6 in tuberculosis pathogenesis and virulence has been confirmed by inactivating *esxA* and *esxB* in virulent *M. bovis* and *M. tuberculosis*, which led to attenuated strains [5,8], and through the reintroduction of RD1 back into BCG Pasteur, which led to increased virulence of the recombinant strain [3].

Neither CFP-10 nor ESAT-6 show any significant sequence similarity with any proteins of known tertiary structure or function. However, they are members of a

large family of mycobacterial proteins, including 23 found in *M. tuberculosis* (Fig. 1) [9], which as with CFP-10 and ESAT-6 are almost always encoded by genes arranged in pairs in the genome. Analysis of the recently published *M. leprae* genome revealed an extreme case of reductive genomic evolution, which has been proposed to define the minimal gene set for a pathogenic mycobacterium [10]. Homology searches for CFP-10/ESAT-6 family members from *M. tuberculosis* against the predicted protein expression profile of *M. leprae* revealed that CFP-10 and ESAT-6 are individually conserved within the *M. leprae* genome (ML0050 and ML0049, respectively). The retention of both CFP-10 and ESAT-6 as functional genes in *M. leprae* clearly demonstrates their importance in the lifestyle of mycobacterial pathogens.

Several other CFP-10/ESAT-6 family members are known to be secreted and are immunogenic, including the products of Rv0287/Rv0288 (*esxG/esxH*), Rv3019c (*esxR*) and members of the pairing groups shown as A

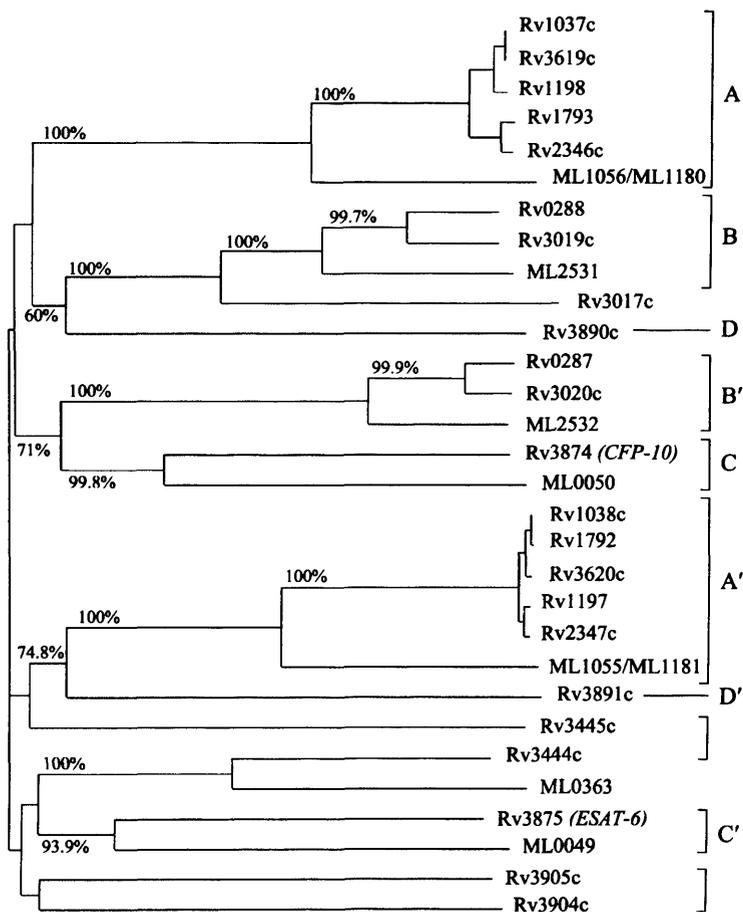


Fig. 1. Phylogenetic tree for the CFP-10/ESAT-6 family of *M. tuberculosis* proteins (prefixed by Rv) and their *M. leprae* homologues (prefixed by ML). Major pairing groups are highlighted by square brackets and labeled (A pairs with A' etc.). Bootstrap values (%) are indicated for the major branch points in the tree. To generate the family relationships shown in the tree, the protein sequences were aligned using the Blosum62 matrix with a gap opening penalty of 12 and a gap extension penalty of 2 per residue.

and A' in Fig. 1 [11,12]. We have recently reported that CFP-10 and ESAT-6 form a tight, 1:1 complex (20.6 kDa), which induces the folding of both proteins and clearly represents their functional form [9]. In our previous communication we proposed that other genome paired members of the CFP-10/ESAT-6 family would form similar, tight, 1:1 heterodimer complexes, and also raised the possibility of complex formation between non-gene partners, which could provide greater functional diversity. In this communication we report the results of comprehensive yeast two-hybrid and biochemical characterisation of complex formation between six members of the family (CFP-10, ESAT-6 and the products of the genes Rv0287/Rv0288 (*esxG/esxH*) and Rv3019c/Rv3020c (*esxR/esxS*)), which provide guidelines to predict complex formation amongst this important protein family.

2. Materials and methods

2.1. Yeast two-hybrid strains and growth conditions

Escherichia coli DH5 α cells transformed with yeast two-hybrid expression vectors were grown at 37 °C in Luria–Bertani (LB) medium supplemented with 100 μ g/ml ampicillin. The yeast two-hybrid assays were carried out in *Saccharomyces cerevisiae* CG1945 cells, which were grown at 30 °C on either YPD (yeast extract/peptone/dextrose) or on minimal dropout media, as appropriate [13]. The yeast two-hybrid vectors carry the nutritional markers *LEU2* (pGAD-C1) and *TRP1* (pGBD-C1), which allows yeast cells transformed with these plasmids to grow on minimal media lacking leucine or tryptophan.

2.2. Expression and purification of Rv0287 and Rv0288

The full-length coding regions for Rv0287 (*esxG*) and Rv0288 (*esxH*) were PCR amplified from the bacterial artificial chromosome (BAC) Rv167 [14] and ligated into a pET28a (Novagen) expression vector. This generated constructs encoding an N-terminal His-tag and thrombin cleavage site (MGSSHHHHHSGLVPR-GSH) followed by the protein coding sequence. Expression was carried out in *E. coli* BL21 (DE3) cells at 37 °C, which were induced for 3 h by the addition of isopropyl β -D-thiogalactopyranoside (IPTG) (0.45 mM) at an optical density of 0.6 at 600 nm. Under these expression conditions both Rv0287 and Rv0288 were found to be insoluble and were initially isolated as inclusion bodies [9]. Rv0287 and Rv0288 were refolded by solubilising the inclusion body pellets in a 25 mM NaH₂PO₄, 6 M guanidine HCl buffer (pH 7.4) at a protein concentration of 0.5 mg/ml, followed by dialysis against the same buffer without the denaturant at 4 °C. The soluble pro-

tein obtained was purified by nickel affinity chromatography. Samples were loaded onto Chelating Sepharose FF (Pharmacia) pre-treated with NiSO₄ and equilibrated in 25 mM NaH₂PO₄, 0.2 M NaCl, 30 mM imidazole buffer (pH 7.4). Rv0287 and Rv0288 were eluted from the column in the sodium phosphate buffer containing 0.5 M imidazole.

2.3. Intrinsic fluorescence

Samples of Rv0287 and Rv0288 (1 μ M) were analysed in 20 mM Tris–HCl, 85 mM NaCl buffer (pH 8.0) at 25 °C using a Varian Eclipse fluorimeter, with the excitation wavelength set to 280 nm. The denaturation curve for the Rv0287/Rv0288 complex was similarly obtained but with samples containing increasing concentrations of guanidine HCl and allowed to equilibrate for 30 min at 25 °C before analysis.

2.4. Construction of pGAD-C1 and pGBD-C1 vectors

The full length coding sequences for CFP-10 (Rv3874, *esxB*) and ESAT-6 (Rv3875, *esxA*) were amplified by PCR from the BAC Rv414 [14]. The purified PCR products were then cloned into the *Eco*RI and *Bam*HI sites of the *GAL4* transactivation domain (TAD) vector pGAD-C1, and into the *GAL4* DNA binding domain (DBD) vector pGBD-C1 [15]. The same approach was also used to produce pGAD-C1 and pGBD-C1 vectors containing Rv0287 (*esxG*) and Rv0288 (*esxH*) from BAC Rv167 and Rv3019c (*esxR*) and Rv3020c (*esxS*) from BAC Rv280 [14].

2.5. Yeast two-hybrid assays

All transformations of *S. cerevisiae* cells with pGAD-C1 and pGBD-C1 vectors were performed using the lithium acetate, single-stranded carrier DNA, polyethylene glycol method [13]. Yeast singularly transformed with the series of pGBD-C1 constructs alone (pGBD-ESAT-6, pGBD-CFP-10, pGBD-Rv0287, pGBD-Rv0288, pGBD-Rv3019c and pGBD-Rv3020c) were selected by growth on minimal dropout media lacking tryptophan and subsequently tested for auto-activation of the *lacZ* reporter gene by colony filter lift assays, using 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal) as the substrate [16]. Yeast cells successfully transformed with specific combinations of pGAD and pGBD vectors were identified by their ability to grow on yeast minimal dropout media lacking leucine and tryptophan. The activity of the *lacZ* reporter gene (indicative of complex formation) in doubly transformed yeast cells was determined by both colony filter lift (X-gal) assays and by quantitative β -galactosidase assays of soluble cell extracts, using *o*-nitrophenyl- β -galactoside (ONPG) as the substrate [16].

3. Results

3.1. Complex formation between genome partners of the CFP-10/ESAT-6 family

We have previously reported that the genome partners CFP-10 and ESAT-6 form a tight, 1:1 heterodimeric complex [9]. We have used similar *in vitro* techniques to demonstrate that Rv0287 and Rv0288 also form a stable complex. Rv0287 and Rv0288 were expressed as His-tagged proteins in *E. coli* where they were found to be insoluble and recovered as inclusion bodies. A simple refolding protocol was used to generate soluble protein that was subsequently purified by nickel-affinity chromatography. Electrospray mass spectrometry confirmed the authenticity of the proteins produced and revealed that the N-terminal Met residue of Rv0287 protein was cleaved, whereas the N-terminal Met residue from Rv0288 was not.

Rv0288 contains 4 tryptophan residues whereas Rv0287 contains none; therefore intrinsic fluorescence provides a convenient method to investigate complex formation. Fig. 2(a) shows fluorescence scans for Rv0287 and Rv0288, both alone and when present together in equimolar concentrations. The wavelength of maximum emission (λ_{\max}) for Rv0288 is 351 nm, con-

sistent with the tryptophan residues of this protein being in an exposed polar environment. For the 1:1 mixture of Rv0287 and Rv0288, the λ_{\max} is blue-shifted to 342 nm indicating the formation of a complex where at least one of the tryptophan residues is moved to a more hydrophobic environment.

A denaturation curve for the Rv0287/Rv0288 complex was constructed by monitoring the change in λ_{\max} on increasing the concentration of guanidine hydrochloride (Fig. 2(b)). The complex showed the properties expected for a folded protein, with a region stable to denaturation followed by a cooperative unfolding transition. The Rv0287/Rv0288 complex is stable to at least 1 M guanidine hydrochloride (GdmCl) and has a midpoint of denaturation at around 2.3 M GdmCl. A similar curve was seen for the CFP-10/ESAT-6 complex [9], except that the Rv0287/Rv0288 complex is more stable with the CFP-10/ESAT-6 complex starting to unfold at ~0.5 M GdmCl and showing a midpoint of denaturation at 0.9 M GdmCl.

Preliminary yeast two-hybrid experiments were designed to determine whether this approach would detect the tight interactions seen *in vitro* for CFP-10/ESAT-6 and Rv0287/Rv0288, and therefore test its suitability to detect interactions between other members of the CFP-10/ESAT-6 family. Yeast cells transformed with combinations of vectors containing the TAD and DBD of *GAL4* fused to CFP-10 and ESAT-6 were assayed for expression of the *lacZ* reporter gene by filter lift based assays. The formation of functional *GAL4* fusion protein complexes, indicated by the development of blue colonies, was only observed for yeast cells co-expressing either TAD-CFP-10 and DBD-ESAT-6 or TAD-ESAT-6 and DBD-CFP-10 (Fig. 3). All other co-transformants remained white, indicating that neither CFP-10 nor ESAT-6 form homodimers and that neither CFP-10 nor ESAT-6 fused to the *GAL4* DBD mediated auto-activation of the reporter gene. These results were confirmed by quantitative β -galactosidase assays shown in Fig. 4(a) and clearly show that the yeast two-hybrid approach is well suited to studying complex formation amongst members of the CFP-10/ESAT-6 family.

Tight interactions between Rv0287 and Rv0288 were identified by both filter lift (results not shown) and quantitative β -galactosidase assays (Fig. 4(b)), but again there was no evidence for the formation of homodimers. Quantitative β -galactosidase assays showed a low level of reporter gene activity in yeast transformed with the DBD-Rv0288 construct and either TAD or TAD-Rv0288, suggesting auto-activation of the *lacZ* reporter gene by the *GAL4* DBD-Rv0288 fusion protein; however, co-expression of DBD-Rv0288 and TAD-Rv0287 resulted in a much higher level of β -galactosidase activity, indicative of complex formation. Reciprocal assays with Rv0287 fused to the *GAL4* DBD only revealed re-

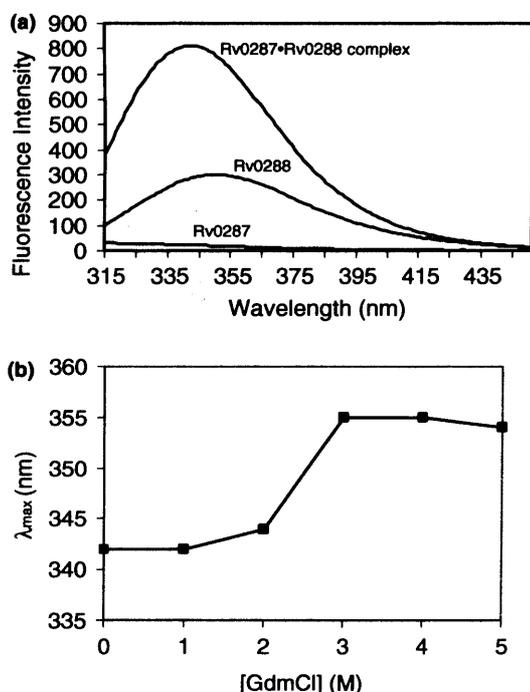


Fig. 2. Complex formation between Rv0287 and Rv0288. (a) Fluorescence scans of equimolar amounts of Rv0287, Rv0288 and the Rv0287/Rv0288 complex (excitation wavelength 280 nm). (b) Guanidine denaturation curve for the Rv0287/Rv0288 complex followed by the shift in wavelength of maximum emission (λ_{\max}).

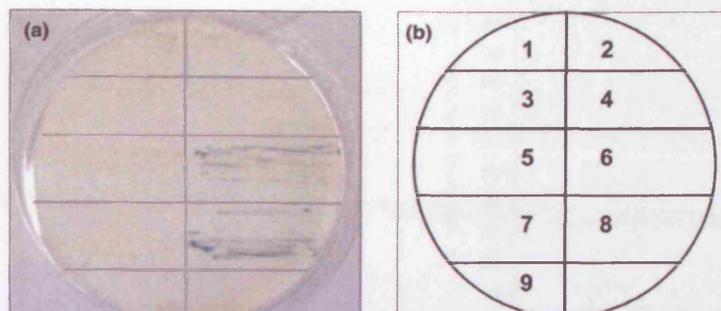


Fig. 3. Yeast two-hybrid detection of pair wise interactions between CFP-10 and ESAT-6. (a) Filter lift assay of yeast co-transformed with pGAD-C1 and pGBD-C1 constructs containing combinations of CFP-10 and ESAT-6 fused to either the *GAL4* TAD or the *GAL4* DBD. Expression of the *lacZ* reporter gene, indicative of complex formation, is detected by the formation of blue colonies. (b) Identification of TAD and DBD fusions used to transform *S. cerevisiae*. 1. TAD + DBD; 2. TAD-ESAT-6 + DBD; 3. TAD-CFP-10 + DBD; 4. TAD + DBD-ESAT-6; 5. TAD-ESAT-6 + DBD-ESAT-6; 6. TAD-CFP-10+DBD-ESAT-6; 7. TAD + DBD-CFP-10; 8. TAD-ESAT-6 + DBD-CFP-10; 9. TAD-CFP-10 + DBD-CFP-10.

porter gene activity when co-expressed with TAD-Rv0288, confirming the pair wise interaction between Rv0287 and Rv0288.

Interaction between Rv3019c and Rv3020c was only detected using yeast doubly transformed with DBD-Rv3019c and TAD-Rv3020c (Fig. 4(c)) as it proved impossible to obtain cells transformed with vector combinations that included DBD-Rv3020c. Transformation with the DBD-Rv3019c vector alone induced some expression of the *lacZ* reporter gene, however significantly increased expression was seen in yeast co-transformed with DBD-Rv3019c and TAD-Rv3020c (Fig. 4(c)), indicating that like CFP-10/ESAT-6 and Rv0287/Rv0288, Rv3019c and Rv3020c interact with each other to form a heterodimer.

3.2. Characterisation of interactions between non-genome partners of the CFP-10/ESAT-6 family

Yeast two-hybrid experiments were employed to determine whether CFP-10/ESAT-6 family proteins are capable of forming tight complexes with members other than their genome partner. CFP-10 and ESAT-6 fused to the *GAL4* TAD or DBD were screened against all possible combinations of constructs encoding Rv0287, Rv0288, Rv3019c and Rv3020c, as shown in Table 1. Results from filter lift assays (not shown) for CFP-10 and ESAT-6 provided no evidence for the formation of functional complexes with any of the other members of the CFP-10/ESAT-6 family tested.

The results of quantitative β -galactosidase assays for yeast cells transformed with Rv3019c fused to the TAD and CFP-10, ESAT-6, Rv0287 or Rv0288 attached to the DBD of *GAL4*, and vice versa, are shown in Fig. 5(a). The data supports the results from filter lift assays confirming that Rv3019c is not able to interact with either CFP-10 or ESAT-6. The low level of activity observed with TAD-Rv3019c/DBD-Rv0288, in both filter lift and quantitative assays, is due to auto-activation

by the DBD-Rv0288 protein as described previously. However, the high β -galactosidase activity for the Rv3019c/Rv0287 combination (Fig. 5(a)) suggests a tight interaction between these fusion proteins. This is confirmed by the data observed for the DBD-Rv3019c construct screened against the TAD vectors for CFP-10, ESAT-6, Rv0287 and Rv0288. In the data reporting interactions between genome pairs there is clear evidence of auto-activation of the reporter gene in cells transformed with the DBD-Rv3019c construct, however, the reporter gene activity was clearly much higher in cells co-transformed with TAD-Rv0287, which confirms that Rv3019c and Rv0287 form a tight complex (Fig. 5(a)).

Interactions involving Rv3020c could only be investigated in yeast cells expressing the TAD-Rv3020c construct, as it was not possible to transform yeast cells with combinations including Rv3020c fused to the DBD of *GAL4*. Both filter lift assays (results not shown) and quantitative β -galactosidase assays (Fig. 5(b)), detected no expression of the *lacZ* reporter gene in yeast cells co-transformed with the TAD-Rv3020c construct in combination with the *GAL4* DBD fused to either CFP-10, ESAT-6 or Rv0287. However, the β -galactosidase activity recorded for the TAD-Rv3020c and DBD-Rv0288 combination is clearly above the background level of auto-activation seen with the TAD-Rv3019c and DBD-Rv0288 combination (Fig. 5(a)), indicating that Rv3020c and Rv0288 are also capable of forming a heterodimeric complexes.

4. Discussion

The functional significance of the CFP-10/ESAT-6 gene cluster to tuberculosis pathogenesis has been well established. However, the determination of the molecular activity and mechanisms of action of CFP-10/ESAT-6 in tuberculosis pathogenesis remains elusive and is a

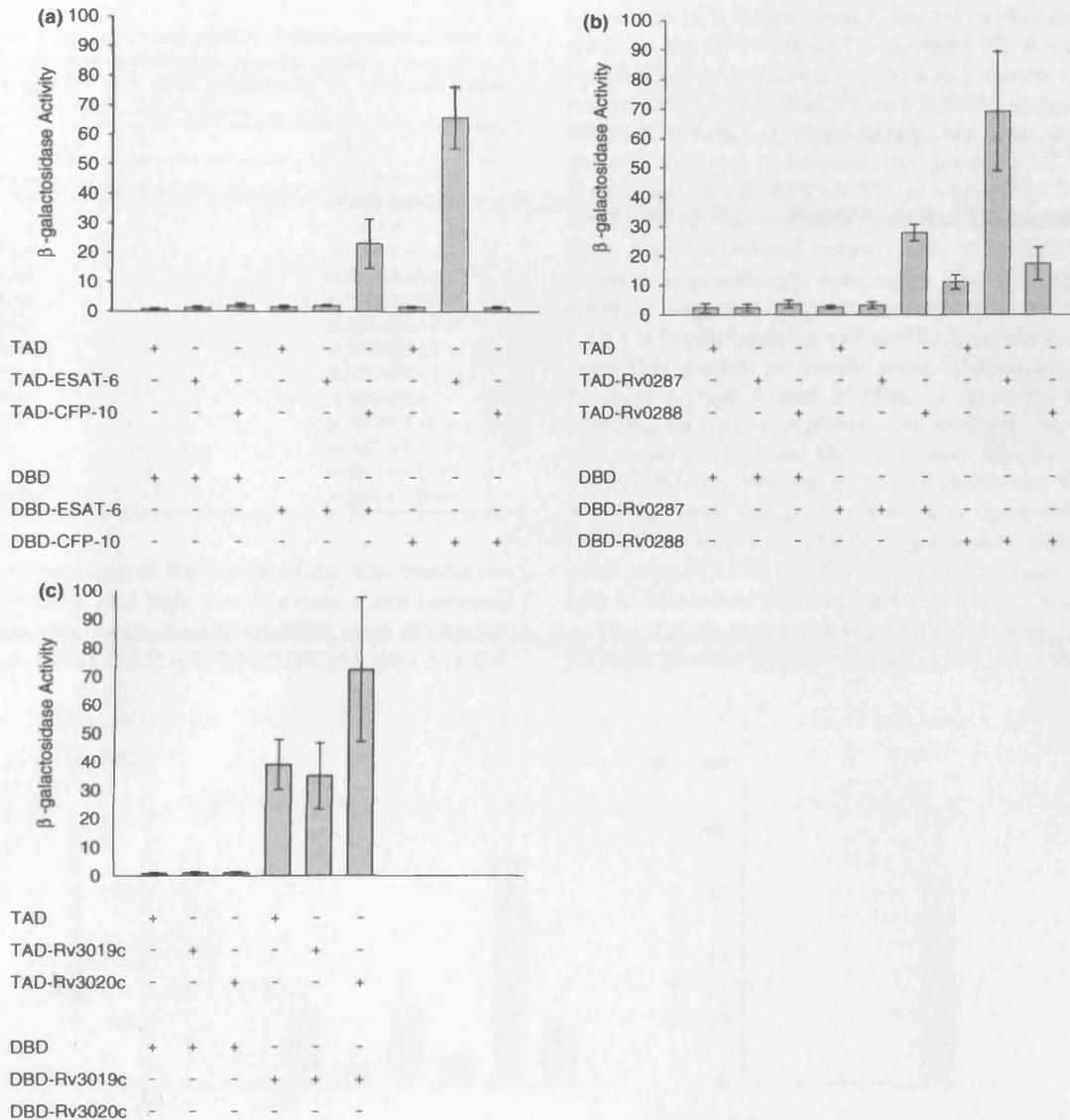


Fig. 4. Complex formation between genome partners of the CFP-10/ESAT-6 family. The graphs shown report mean values from four independent assays of β -galactosidase activity (nmol/mg/min) in cell-free extracts obtained from yeast cells co-expressing pGAD-C1 and pGBD-C1 constructs encoding (a) ESAT-6 and CFP-10 (b) Rv0287 and Rv0288 and (c) Rv3019c and Rv3020c. The results clearly show complex formation between the paired genome products.

major outstanding challenge. We have recently reported that CFP-10 and ESAT-6 form a tight 1:1 complex, which induces the folding of both proteins and probably represents their functional form [9]. There are a further 20 members of the CFP-10/ESAT-6 family arranged in pairs within the TB genome and we previously suggested that these paired family members would also be co-ordinately regulated, with the paired protein products forming similar, tight 1:1 complexes. In addition, we raised the question of whether complex formation of CFP-10/ESAT-6 family members was limited only to their gene partners, or possibly more widespread leading to many

more than 11 functional protein complexes, thereby enhancing the functional flexibility of the CFP-10/ESAT-6 family of proteins, which could be important for virulence of members of the *M. tuberculosis* complex.

The work reported here on characterisation of complex formation between members of the CFP-10/ESAT-6 protein family has concentrated on CFP-10/ESAT-6 and the products of Rv0287/Rv0288 (*esxG/esxH*), and the phylogenetically related pair Rv3019c/Rv3020c (*esxR/esxS*) (over 95% amino acid sequence homology with their respective Rv0288/Rv0287 counterparts), which were considered likely candidates to inter-

Table 1
Combinations of pGAD-C1 and pGBD-C1-based constructs used to transform *S. cerevisiae* to investigate possible complex formation of CFP-10 and ESAT-6 with other members of the CFP-10/ESAT-6 family

pGAD-C1 Construct	pGBD-C1 Construct
pGAD-ESAT-6	pGBD-Rv0287
pGAD-ESAT-6	pGBD-Rv0288
pGAD-ESAT-6	pGBD-Rv3019c
pGAD-CFP-10	pGBD-Rv0287
pGAD-CFP-10	pGBD-Rv0288
pGAD-CFP-10	pGBD-Rv3019c
pGAD-Rv0287	pGBD-ESAT-6
pGAD-Rv0288	pGBD-ESAT-6
pGAD-Rv3019c	pGBD-ESAT-6
pGAD-Rv3020c	pGBD-ESAT-6
pGAD-Rv0287	pGBD-CFP-10
pGAD-Rv0288	pGBD-CFP-10
pGAD-Rv3019c	pGBD-CFP-10
pGAD-Rv3020c	pGBD-CFP-10

act with the products of Rv0287/Rv0288. The results reported here show that tight binding does occur between the products of both Rv0287/Rv0288 and Rv3019c/Rv3020c and that the Rv0287/Rv0288 complex has the

properties of a folded protein similar to that previously seen for the CFP-10/ESAT-6 complex [9]. Recent work by Okkels and Andersen [17] has also shown a binding interaction between Rv0287 and Rv0288 using Western blotting techniques. Importantly, we also show that there is cross talk between the products of Rv0287/Rv0288 and Rv3019c/Rv3020c, in which the CFP-10 related gene products (Rv0287 and Rv3020c) interact with their ESAT-6 related counterparts (Rv0288/Rv3019c) forming heterodimeric complexes. Together, this work strongly suggests that all genome pairs of CFP-10/ESAT-6 family proteins will bind to each other and indicates that groups of closely related sequences, such as amongst groups A and A' (Fig. 1), probably have the potential to form complexes with non-genome partners and therefore increase the functional flexibility of this protein family. Neither Rv0287/Rv0288 nor Rv3019c/Rv3020c were found to show any interaction with CFP-10 or ESAT-6, which suggests that distantly related pairs of CFP-10/ESAT-6 family proteins are unlikely to interact with other members of the family.

The closely related Rv0287/Rv0288 and Rv3019c/Rv3020c genome pairs are conserved as just a single pair

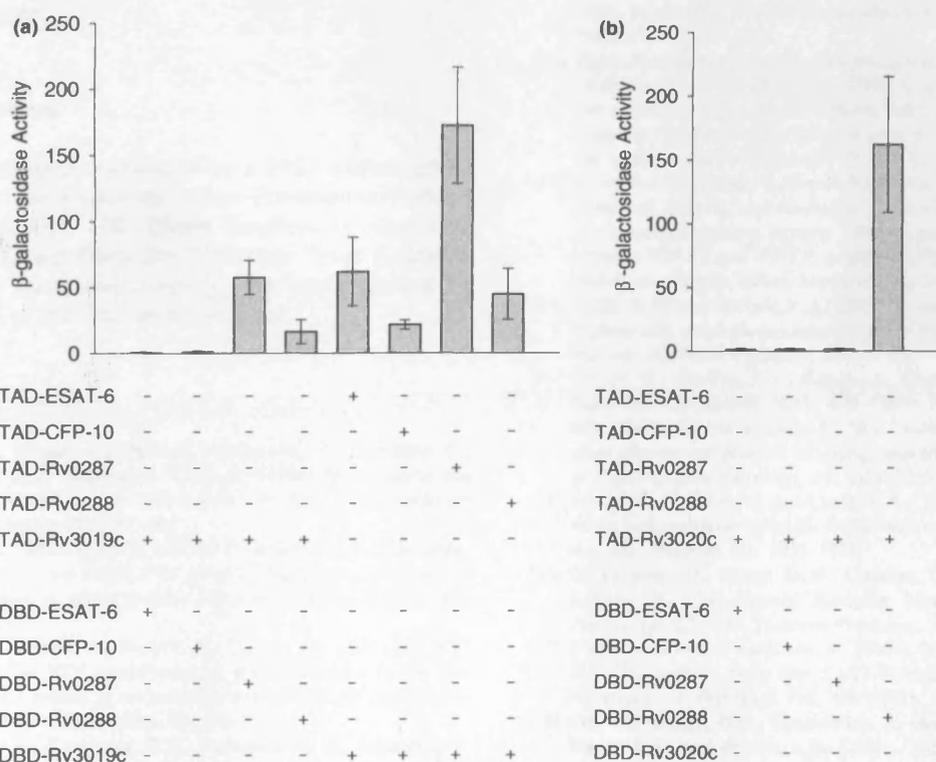


Fig. 5. Interactions between non-genome partners of the CFP-10/ESAT-6 family. The histograms shown report the results of four independent β -galactosidase activity assays (nmol/mg/min) of soluble cell extracts from *S. cerevisiae* co-transformed with various combinations of pGAD-C1 and pGBD-C1 constructs. (a) Analysis of possible interactions between Rv3019c fused to the *GALA* TAD and the *GALA* DBD fused to either CFP-10, ESAT-6, Rv0287 or Rv0288, and vice versa. (b) Investigation of complex formation between Rv3020c fused to the *GALA* TAD and CFP-10, ESAT-6, Rv0287 and Rv0288 fused to the DBD of *GALA*.

in the genome of *M. leprae* (Fig. 1) which suggests that they may be functionally redundant in the case of *M. tuberculosis* complex cells. Furthermore, the cluster of genes including Rv3019c/Rv3020c lacks several homologues of the genes neighbouring CFP-10/ESAT-6 which appear to code for a protein complex responsible for the active secretion of CFP-10/ESAT-6 family proteins [6,7,18]. The Rv0287/Rv0288 gene cluster 3 [18], however, contains homologues of all the proposed components of the secretory complex, as does the conserved gene cluster within the *M. leprae* genome. The product of the Rv3019c gene has been detected in culture supernatants, which perhaps suggests that the close sequence similarity between Rv0287/Rv0288 and Rv3019c/Rv3020c allows secretion of the Rv3019c/Rv3020c complex via the complex formed from components of the Rv0287/Rv0288 gene cluster.

In conclusion, the data reported here clearly indicate the possibility for some promiscuity in complex formation between closely sequence-related members of the CFP-10/ESAT-6 protein family, which may provide greater functional flexibility. However, members of the family that are distantly related from other members, such as CFP-10/ESAT-6 and Rv3904c/Rv3905c, are probably constrained to form complexes with only their genome partner.

Acknowledgement

Kirsty Lightbody is funded by a PhD student grant awarded by the Department for Environment, Food and Rural Affairs, UK. Philip Renshaw is supported by a project grant from the Wellcome Trust (066047). Electrospray mass spectrometry data was collected by K. Howland at the University of Kent.

References

- [1] Cole, S.T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D. and Gordon, S.V., et al. (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544.
- [2] Behr, M.A., Wilson, M.A., Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S. and Small, P.M. (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284, 1520–1523.
- [3] Pym, A.S., Brodin, P., Brosch, R., Huerre, M. and Cole, S.T. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol. Microbiol.* 46, 709–717.
- [4] Berthet, F.-X., Ramussen, P.B., Rosenkrands, I., Andersen, P. and Gicquel, B. (1998) A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). *Microbiology* 144, 3195–3203.
- [5] Hsu, T., Hinigley-Wilson, S.M., Chen, B., Chen, M., Dai, A.Z., Morin, P.M., Marks, C.B., Padiyar, J., Goulding, C., Gingery, M., Eisenberg, D., Russell, R.G., Derrick, S.C., Collins, F.M., Morris, S.L., King, C.H. and Jacobs Jr., W.R. (2003) The primary mechanism of attenuation of bacillus Calmette–Guérin is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc. Natl. Acad. Sci. USA* 100, 12420–12425.
- [6] Stanley, S.A., Raghaven, S., Hwang, W.H. and Cox, J.S. (2003) Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialised secretion system. *Proc. Natl. Acad. Sci. USA* 100, 13001–13006.
- [7] Guinn, K.M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., Smith, S. and Sherman, D.R. (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol. Microbiol.* 51, 359–370.
- [8] Wards, B.J., de Lisle, G.W. and Collins, D.M. (2000) An *esat6* knockout mutant of *Mycobacterium bovis* produced by recombination will contribute to the development of a live tuberculosis vaccine. *Tubercle Lung Dis.* 80, 185–189.
- [9] Renshaw, P.S., Panagiotidou, P., Whelan, A., Gordon, S.V., Hewinson, R.G., Williamson, R.A. and Carr, M.D. (2002) Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterisation of the structural properties of ESAT-6, CFP-10 and the ESAT-6/CFP-10 complex: Implications for pathogenesis and virulence. *J. Biol. Chem.* 277, 21598–21603.
- [10] Cole, S.T., Eiglmeier, K., Parkhill, J., James, K.D. and Thomson, N.R., et al. (2001) Massive gene decay in the Leprosy Bacillus. *Nature* 409, 1007–1011.
- [11] Skjöt, R.L.V., Oettinger, T., Rosenkrands, I., Ravn, P., Brock, I., Jacobsen, S. and Andersen, P. (2000) Comparative evaluation of low-molecular-mass proteins from *Mycobacterium tuberculosis* identifies members of the ESAT-6 family as immunodominant T-cell antigens. *Infect. Immunol.* 68, 214–220.
- [12] Skjöt, R.L.V., Brock, I., Arend, S.M., Munk, M.E., Theisen, M., Ottenhoff, T.H.M. and Andersen, P. (2002) Epitope mapping of the immunodominant antigen TB10.4 and the two homologous proteins TB10.3 and Tb12.9, which constitute a subfamily of the *esat-6* gene family. *Infect. Immunol.* 70, 5446–5453.
- [13] Gietz, R.D. and Woods, R.A. (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* 350, 87–96.
- [14] Brosch, R., Gordon, S.V., Billault, A., Garnier, T., Eiglmeier, K., Soravito, C., Barrell, B.G. and Cole, S.T. (1998) Use of a *Mycobacterium tuberculosis* H37Rv bacterial artificial chromosome library for genome mapping, sequencing and comparative genomics. *Infect. Immunol.* 66, 2221–2229.
- [15] James, P., Halliday, J. and Craig, E.A. (1996) Genomic libraries and a host strain designed for highly efficient two hybrid selection in yeast. *Genetics* 144, 1425–1436.
- [16] Le Douarin, B., Heery, D.M., Gaudon, C., vom Baur, E. and Losson, R. (2001) Steroid Receptor Methods: Protocols and Assays, pp. 227–248. Humana Press Inc., Totowa, NJ.
- [17] Okkels, L.M. and Andersen, P. (2004) Protein–protein interactions of proteins from the ESAT-6 family of *Mycobacterium tuberculosis*. *J. Bacteriol.* 186, 2487–2491.
- [18] Gey van Pittius, N.C., Gamielien, J., Hide, W., Brown, G.D., Siezen, R.J., and Beyers, A.D. (2001). The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G + C gram-positive bacteria. *Genome Biol.* 2(10): Research, 0044.1-0044.18.

Structure and function of the complex formed by the tuberculosis virulence factors CFP-10 and ESAT-6

Philip S Renshaw¹, Kirsty L Lightbody¹,
Vaclav Veverka¹, Fred W Muskett¹,
Geoff Kelly², Thomas A Frenkiel²,
Stephen V Gordon³, R Glyn Hewinson³,
Bernard Burke⁴, Jim Norman¹,
Richard A Williamson⁵ and Mark D Carr^{1,*}

¹Department of Biochemistry, University of Leicester, Leicester, UK, ²MRC Biomedical NMR Centre, National Institute for Medical Research, Mill Hill, London, UK, ³TB Research Group, Veterinary Laboratories Agency, Addlestone, Surrey, UK, ⁴Department of Infection, Immunity and Inflammation, University of Leicester, Leicester, UK and ⁵Department of Biosciences, University of Kent, Canterbury, Kent, UK

The secreted *Mycobacterium tuberculosis* complex proteins CFP-10 and ESAT-6 have recently been shown to play an essential role in tuberculosis pathogenesis. We have determined the solution structure of the tight, 1:1 complex formed by CFP-10 and ESAT-6, and employed fluorescence microscopy to demonstrate specific binding of the complex to the surface of macrophage and monocyte cells. A striking feature of the complex is the long flexible arm formed by the C-terminus of CFP-10, which was found to be essential for binding to the surface of cells. The surface features of the CFP-10·ESAT-6 complex, together with observed binding to specific host cells, strongly suggest a key signalling role for the complex, in which binding to cell surface receptors leads to modulation of host cell behaviour to the advantage of the pathogen.

The EMBO Journal (2005) 24, 2491–2498. doi:10.1038/sj.emboj.7600732; Published online 23 June 2005

Subject Categories: structural biology; microbiology & pathogens

Keywords: CFP-10; ESAT-6; pathogenesis; tuberculosis; virulence

Introduction

Tuberculosis kills 2–3 million people annually (World Health Organisation, 2004), yet we still have little understanding of the molecular basis of pathogenesis. Comparisons of the genomes of virulent *Mycobacterium tuberculosis* and *Mycobacterium bovis* with attenuated BCG vaccine strains identified a deletion (RD1) in all BCG strains that plays a key role in virulence (Mahairas *et al.*, 1996; Behr *et al.*, 1999; Brosch *et al.*, 2001; Pym *et al.*, 2002). This region contains nine protein-coding genes (Rv3871–3879c) and subsequent work

showed that inactivation of just two of these (Rv3874 and Rv3875), coding for the secreted proteins CFP-10 (100 residues) and ESAT-6 (95 residues), results in dramatically reduced virulence (Behr *et al.*, 1999; Wards *et al.*, 2000; Pym *et al.*, 2002; Hsu *et al.*, 2003; Stanley *et al.*, 2003). These proteins clearly play an essential role in tuberculosis pathogenesis, but show no homology to any proteins of known structure or function.

CFP-10 and ESAT-6 are members of a large family of mycobacterial proteins, including 22 found in *M. tuberculosis*, which are encoded by genes arranged in pairs in the genome (Renshaw *et al.*, 2002). In the case of CFP-10 and ESAT-6, the genes have been shown to be coordinately regulated (Berthet *et al.*, 1998) and both proteins are secreted despite lacking a conventional signal sequence (Pym *et al.*, 2003). Recent work indicates that secretion of the proteins is an active process involving a membrane protein complex formed from the products of several flanking genes (Gey van Pittius *et al.*, 2001; Hsu *et al.*, 2003; Sasseti and Rubin, 2003; Stanley *et al.*, 2003; Guinn *et al.*, 2004). Several other CFP-10/ESAT-6 family members are also known to be secreted, including the products of Rv0287/Rv0288 and Rv3019c/Rv3020c (Alderson *et al.*, 2000; Rosenkrands *et al.*, 2000; Skjølø *et al.*, 2000, 2002), which we have recently shown form tight complexes, as observed for CFP-10 and ESAT-6 (Renshaw *et al.*, 2002; Lightbody *et al.*, 2004). The *Mycobacterium leprae* genome contains only 1604 functional protein genes compared to 4006 for *M. tuberculosis* and has been proposed to represent the minimal gene set for a pathogenic mycobacterium (Cole *et al.*, 2001). Strikingly, of the 11 pairs of CFP-10/ESAT-6 family proteins found in *M. tuberculosis*, only orthologues of CFP-10 and ESAT-6 are individually conserved in *M. leprae* (ML0050c and ML0049c, respectively), which further emphasises their importance in the lifecycle of mycobacterial pathogens (Renshaw *et al.*, 2002).

Recently, we showed that CFP-10 and ESAT-6 form a tight ($K_d \leq 1.1 \times 10^{-8}$ M), 1:1 complex (Renshaw *et al.*, 2002) and here we report the solution structure of the complex, together with fluorescence microscopy data showing specific binding of the complex to the surface of primary macrophages, the main cell type infected by *M. tuberculosis*. The structural features of the CFP-10·ESAT-6 complex, together with clear evidence for specific binding of the complex to the surface of host cells, strongly imply a signalling role for the CFP-10·ESAT-6 complex, in which binding to cell surface receptors may lead to modulation of host cell behaviour.

Results and discussion

We have determined the solution structure of the CFP-10·ESAT-6 complex to high precision, which is clearly evident from the overlay of the protein backbone shown for the family of 28 satisfactorily converged structures in Figure 1.

*Corresponding author. Department of Biochemistry, University of Leicester, Henry Wellcome Building, Leicester LE1 7HN, UK.
Tel.: +44 116 229 7075; Fax: +44 116 229 7018;
E-mail: mdc12@le.ac.uk

Received: 5 January 2005; accepted: 7 June 2005; published online: 23 June 2005

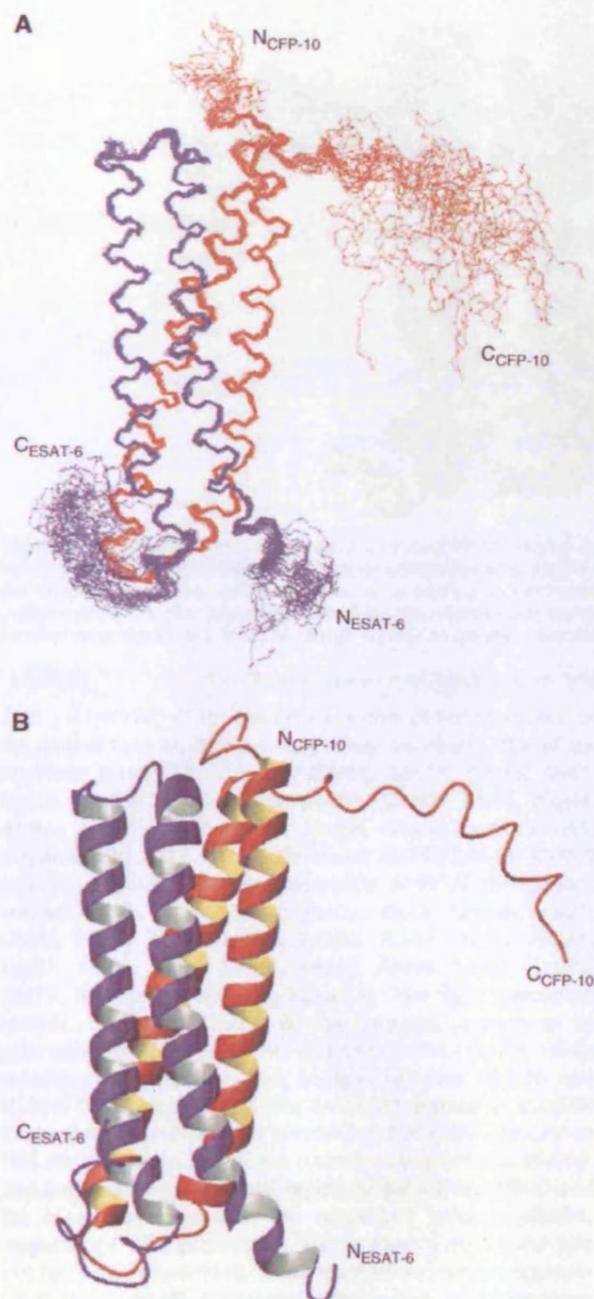


Figure 1 Solution structure of the CFP-10 · ESAT-6 complex. (A) A best-fit superposition of the protein backbone for the family of 28 converged structures obtained, with CFP-10 shown in red and ESAT-6 in blue. The long flexible C-terminal arms of both proteins are clearly visible, as is the propensity to helical structure in this region of CFP-10. (B) A ribbon representation of the backbone topology of the CFP-10 · ESAT-6 complex based on the converged structure closest to the mean, which illustrates the two helix-turn-helix hairpin structures formed by the individual proteins. The orientation of the complex is identical to that shown in panel A, with CFP-10 in red and ESAT-6 in blue. The helical propensity of residues 85–95 in the flexible C-terminus of CFP-10 can be clearly seen in the top right of the figure.

This is also reflected in low root mean squared deviation values to the mean structure for both the backbone and all heavy atoms (0.49 ± 0.13 and 0.93 ± 0.12 Å, respectively) for

the well-defined regions of the complex (residues 6–85 in CFP-10 and ESAT-6). The family of converged structures contain no distance or van der Waals violations greater than 0.5 Å and no dihedral angle violations greater than 5°, with an average value for the CYANA target function of 7.41 ± 0.85 Å² (Herrmann *et al*, 2002). The sums of the violations for upper distance limits, lower distance limits, van der Waals contacts and torsion angle constraints were 34.8 ± 2.44 Å, 2.1 ± 0.28 Å, 13.3 ± 1.29 Å and $81.3 \pm 12.54^\circ$, respectively. Similarly, maximum violations for the converged structures were 0.35 ± 0.05 Å, 0.25 ± 0.03 Å, 0.27 ± 0.06 Å and $3.95 \pm 0.45^\circ$, respectively. Analysis of the backbone dihedral angles for the family of converged structures using PROCHECK (Laskowski *et al*, 1996) revealed that 82% of the residues adopt backbone conformations found within the most favoured regions of a Ramachandran plot and that 17% lie within the additional allowed regions, with no residues consistently found in disallowed regions.

The well-defined core of the CFP-10 · ESAT-6 complex consists of two similar helix-turn-helix hairpin structures formed from the individual proteins, which have an extensive hydrophobic contact surface and lie antiparallel to each other to form a four-helix bundle (Figure 1). A striking feature of the complex is the disordered N- and particularly C-termini of both proteins (residues 2–5 and 86–100 in CFP-10 and 1–3 and 86–95 in ESAT-6), which form long flexible arms at both ends of the four-helix bundle core. The two long helices in the hairpin structures are formed from residues Ala8–Gln40 and Ala47–Ala79 in CFP-10, and from Phe8–Trp43 and Glu49–Ala79 in ESAT-6. The helices in CFP-10 are completely α -helical, whereas in ESAT-6 both long helices terminate with a single turn of 3_{10} helix and ESAT-6 also contains a short 3_{10} helix close to the N-terminus (Gln4–Trp6). Chemical shift and NOE data indicate that part of the exposed C-terminal region of CFP-10 (Arg85–Ser95) has a distinct propensity to adopt a helical conformation, which is clearly evident in Figure 1. This region of CFP-10 may be involved in interactions with a host cell target protein (discussed below), resulting in stabilisation of the helical conformation.

The CFP-10 · ESAT-6 complex has recently been proposed to have host cell lysis activity mediated via the formation of pores in cell membranes (Hsu *et al*, 2003); however, analysis of the electrostatic surface of the complex strongly argues against a pore forming role. The surface of the complex has a very uniform distribution of positive and negative charge, with no hint of a significant hydrophobic patch (Figure 2), which is clearly inconsistent with a membrane spanning pore. In addition, the complex is soluble to over 2 mM in aqueous solution with no sign of aggregation, which is certainly not typical behaviour for a pore forming protein. The surface of the complex is also devoid of any striking acidic or basic patches, which in the latter case suggests that the complex is not involved in interactions with nucleic acids. Similarly, there are no significant clefts in the surface of the structure indicative of an enzyme active site, which suggests a noncatalytic role for the complex. Overall, the surface features of the CFP-10 · ESAT-6 complex seem most consistent with a function based on specific binding to one or more target proteins, perhaps playing a key role in pathogen–host cell signalling.

The extensive contact surface between CFP-10 and ESAT-6 is essentially hydrophobic in nature and comprises about

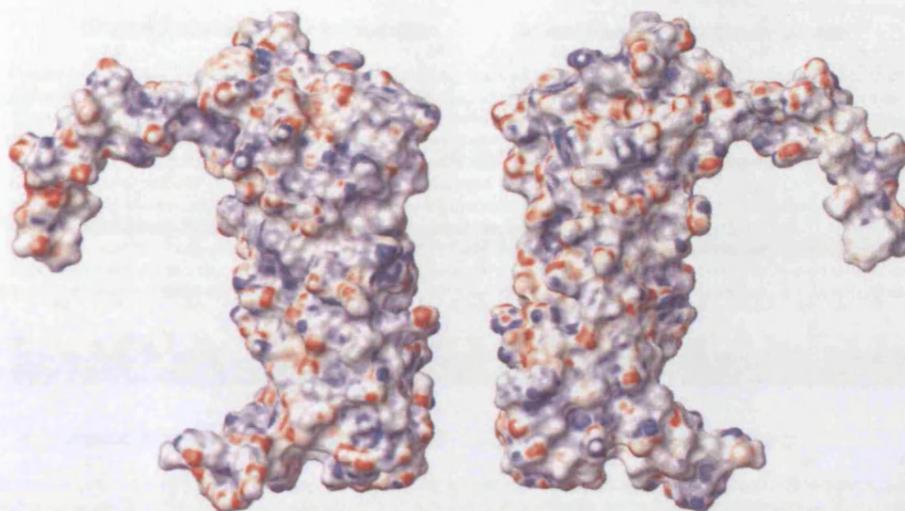


Figure 2 Space-filled views of the surface of the CFP-10·ESAT-6 complex based on the converged structure closest fit to the mean. The right-hand view depicts the complex in the same orientation as in Figure 1, while the left-hand view is rotated by 180° about the vertical axis to show the opposite face. The surface is coloured according to electrostatic potential, with areas of significant negative charge in red, significant positive charge in blue and neutral as white. The electrostatic potential was calculated using MOLMOL (Koradi *et al*, 1996), with the threshold for depicting significant areas of charge chosen to obtain a neutral representation for the fully exposed aromatic ring of Phe100 in CFP-10.

25% (~1800 Å²) of the total surface area of both proteins. In the case of CFP-10, 29 residues account for nearly 90% of the interface (Lys5, Thr6, Leu11, Glu14, Asn17, Phe18, Ile21, Leu25, Gln28, Val32, Thr35, Leu39, Gln42, Trp43, Arg44, Ala46, Ala47, Ala50, Ala54, Phe58, Ala61, Lys64, Gln65, Glu68, Glu71, Ile72, Asn75, Ile76 and Ala79) and for ESAT-6 just 26 residues form approximately 85% of the contact surface (Ile11, Ala14, Ile18, Asn21, Ile25, Leu28, Leu29, Glu31, Gly32, Ser35, Lys38, Leu39, Ala41, Ala42, Trp43, Lys57, Trp58, Thr61, Glu64, Leu65, Ala68, Leu69, Leu72, Thr75, Ile76 and Met83; see Figure 3). The tight interaction between the two proteins in the complex appears to be primarily based on extensive and favourable van der Waals contacts; however, two salt bridges between CFP-10 and ESAT-6 (Glu14–Lys38 and Glu71–Lys57) appear to stabilise interactions between the N-terminal end of helix-1 in CFP-10 and the C-terminal end of the corresponding helix in ESAT-6, and between the C-terminal region of helix-2 in CFP-10 and the N-terminal region of the equivalent helix in ESAT-6, respectively. The positions of the residues forming the two salt bridges are indicated on the multiple sequence alignment for CFP-10- and ESAT-6-related proteins from *M. tuberculosis* shown in Figure 3 and their lack of conservation indicates that these specific salt bridges are not a general feature of the complexes formed by CFP-10/ESAT-6 family proteins, nor a predictor of which family members will form complexes. Analysis of the multiple sequence alignment (Figure 3) reveals that over half of the interface residues are conserved to type in at least two-thirds of the sequences. This, together with predicted helical structures for all members of the *M. tuberculosis* CFP-10/ESAT-6 family and known complex formation for several genome pairs, strongly suggests that all pairs of these proteins will form similar, four-helix bundle containing complexes. However, careful consideration of both the structural and sequence conservation data provides no clear rules for predicting which nongenome paired members of the family will form tight complexes, beyond the

importance of close sequence similarity discussed previously (Lightbody *et al*, 2004).

The multiple sequence alignments shown in Figure 3 reveal that there are a number of hydrophobic and aromatic residues located in the C-terminal regions of CFP-10 (Tyr83 and Leu94) and ESAT-6 (Phe94), which are conserved across the whole family of *M. tuberculosis*-related proteins. These residues are found on the surface of the CFP-10·ESAT-6 complex and play no structural role, implying some functional significance. The C-terminal regions of CFP-10 and ESAT-6 are also as well conserved between *M. tuberculosis* and *M. leprae* as the overall proteins (overall 69% amino-acid sequence homology for CFP-10/ML0050c and 62% for ESAT-6/ML0049c) despite no structural role in the complex, which again implies functional pressure to conserve these regions. It therefore seems possible that the conserved flexible arms in the CFP-10·ESAT-6 complex form part of the interaction site with a target protein.

As discussed above, the structure of the complex strongly suggests that its function is mediated via binding to specific target proteins. The complex has recently been shown to be actively secreted from *M. tuberculosis* and *M. bovis* bacilli (Hsu *et al*, 2003; Pym *et al*, 2003; Stanley *et al*, 2003; Guinn *et al*, 2004), and the expression of both proteins was significantly downregulated by bacteria internalised in macrophages (Schnappinger *et al*, 2003), raising the possibility that target proteins may be found on the surface of host cells. To test this hypothesis, we have covalently labelled the N-termini of both proteins in the complex with a fluorophore (Alexa Fluor 546) and used fluorescence microscopy to look for specific binding to a variety of cell types, including primary monocytes and macrophages, U937 and MonoMac 6 (MM6) monocyte cell lines, and the fibroblast cell lines COS-1 and NIH-3T3.

The primary monocytes and macrophages, together with both monocyte cell lines, consistently showed intense fluorescence at the cell surface after incubation with the labelled

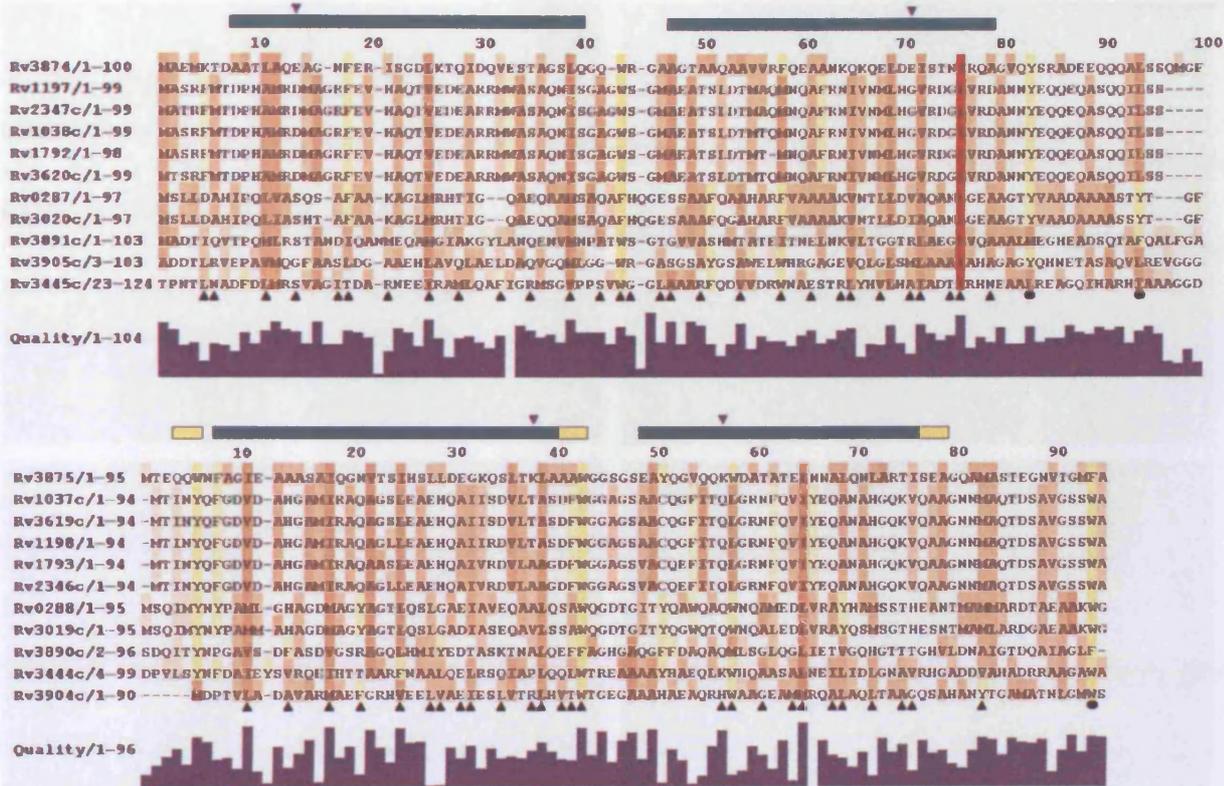


Figure 3 Conservation of the amino-acid sequences of the CFP-10 (top) and ESAT-6 (bottom) related proteins from *M. tuberculosis*. Aliphatic residues with hydrophobic side chains (Leu, Ile, Val, Met and Ala) are highlighted in red and aromatic residues (Phe, Tyr and Trp) in yellow. The extent of sequence conservation is indicated by the histogram quality score and for highlighted residues the greater the conservation, the more intense the colour. Residues forming the interface between CFP-10 and ESAT-6 are indicated by upright black triangular symbols and conserved hydrophobic residues in the C-terminal arms of the proteins by closed circles. Similarly, the locations of four residues forming two intermolecular salt bridges are indicated by down-turned blue triangles. The positions of four residues in the structure of the complex are shown by bars above the sequences (α in green and 3_{10} in yellow). The sequences were aligned using ClustalW, with a standard Blosum30 scoring matrix, a gap opening penalty of 10 and a gap extension penalty of 1 (Thompson *et al.* 1997).

complex, which in a significant proportion of cells was further focused in patches reminiscent of the 'cap-like' structures associated with cell surface receptors (Kwiatkowska and Sobota, 1999). In contrast, no significant fluorescence labelling was seen for the two fibroblast cell lines. This is illustrated by the representative images shown in Figure 4. Analogous experiments were also carried out with Alexa Fluor 546-labelled MPB70 (another major secreted protein of *M. tuberculosis* and *M. bovis*) and MM6 cells. In this case, no significant labelling of the surface of MM6 cells was detected, which strongly argues that the fluorescence localisation observed with the labelled CFP-10·ESAT-6 complex is not mediated by the fluorophore. Similarly, the cell type-specific labelling observed for the CFP-10·ESAT-6 complex suggests that the localisation results from tight binding to a cell surface receptor (probably a protein) expressed in monocytic cells, the main cell type infected by *M. tuberculosis*, rather than a nonspecific interaction between the complex or fluorophore and cell surface. This point was further investigated using U937 cells, which were incubated with fluorescently labelled CFP-10·ESAT-6 complex in the presence of a 20-fold molar excess of unlabelled complex. The intensity of the fluorescence associated with the surface of U937 cells under these conditions was very significantly

reduced (Figure 5), which clearly indicates that binding is mediated by the protein complex and not the attached fluorophore.

In order to test the hypothesis that the binding of the CFP-10·ESAT-6 complex to the surface of target host cells involves the flexible C-termini of either CFP-10 or ESAT-6, complexes formed from truncated CFP-10 (residues 1–86) bound to full-length ESAT-6 and full-length CFP-10 bound to truncated ESAT-6 (residues 1–84) were similarly labelled with Alexa Fluor 546 and incubated with U937 monocytes. The fluorescence labelling observed for cells incubated with the full-length CFP-10·truncated ESAT-6 complex was indistinguishable from that observed for the intact CFP-10·ESAT-6 complex. In contrast, the fluorescence labelling observed for U937 cells incubated with truncated CFP-10 bound to full-length ESAT-6 was dramatically reduced (Figure 6). These findings clearly confirm that binding of fluorescently labelled CFP-10·ESAT-6 complex to the surface of U937 cells is mediated by the protein complex and also show that the flexible C-terminal arm of CFP-10 forms an essential part of the cell surface receptor binding site. It is also worth noting that during time course experiments not reported in detail here, both primary and MM6 cells exposed to the labelled CFP-10·ESAT-6 complex for periods of at least several hours

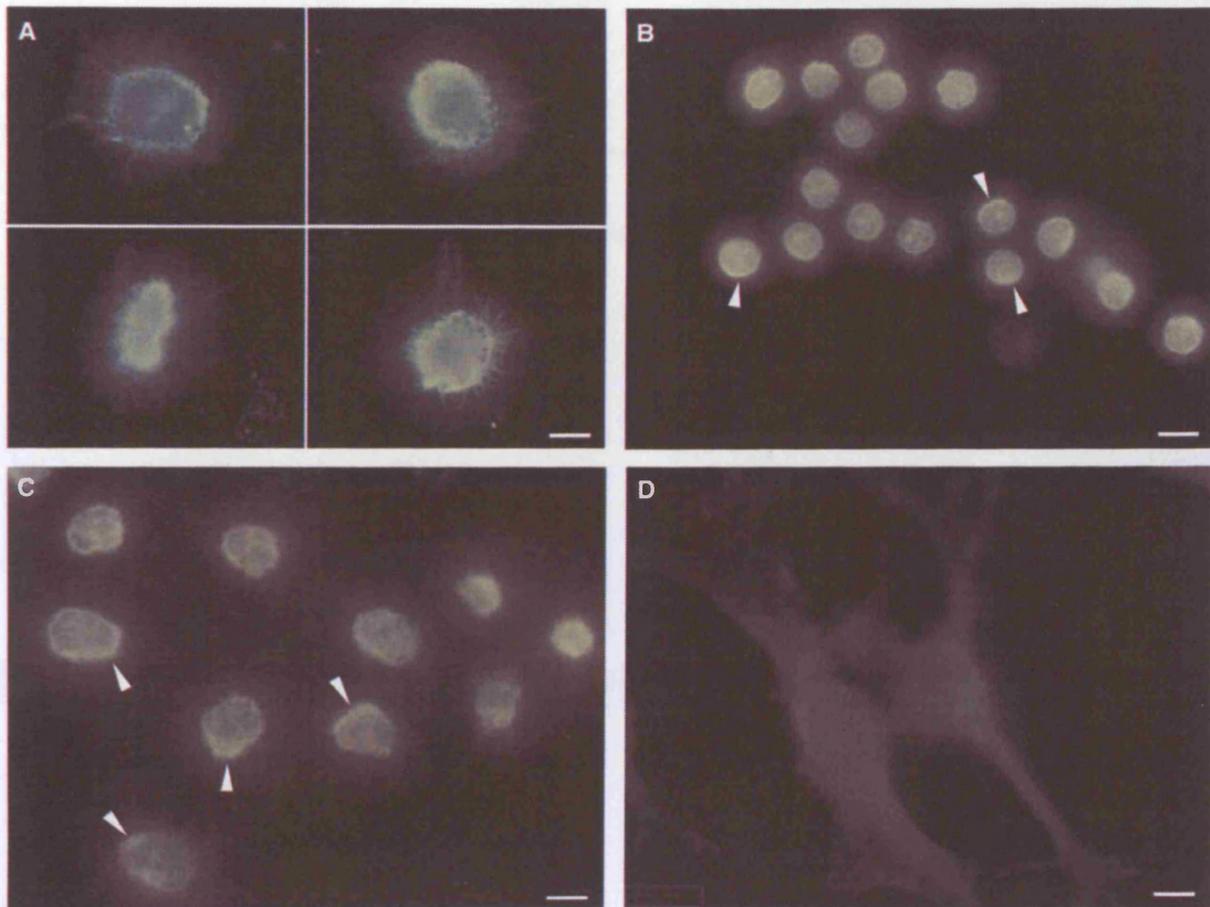


Figure 4 Binding of fluorescently labelled CFP-10·ESAT-6 complex to the surface of monocyte lineage cells. (A–C) (A, C, 100 ms exposure; B, 200 ms exposure) The typical fluorescence localisation observed after incubating primary macrophages, primary monocytes and the monocyte cell line U937 with 1 μ M fluorescently labelled CFP-10·ESAT-6. The fluorescence labelling is focused at the cell surface for all three monocytic cells and is often seen concentrated in patches (indicated by arrows), resembling the 'cap-like' structures associated with receptor-mediated signalling. In contrast, no significant fluorescence labelling was detected for several fibroblast cell lines, as illustrated by the representative images shown for NIH3T3 cells in (D) (200 ms exposure). The size bars shown correspond to 5 μ m.

showed no evidence of lysis, supporting the conclusion that the complex is not associated with cytolytic activity.

The work reported here implies a possible signalling role for the CFP-10·ESAT-6 complex, in which binding to cell surface receptors leads to modulation of host cell behaviour, and clearly represents a major advance in our understanding of the essential role of the CFP-10·ESAT-6 complex in tuberculosis pathogenesis. During final preparation of this manuscript, it was reported that secreted RD1 virulence determinants are required for macrophage aggregation and the subsequent formation of granulomas in zebrafish infected with *Mycobacterium marinum* (Volkman *et al*, 2004). This finding clearly supports our conclusion that the CFP-10·ESAT-6 complex acts as a signalling molecule and further work is now ongoing to identify the host cell target proteins for the complex.

Materials and methods

Protein preparation

The nonlabelled and uniformly ^{15}N - and $^{15}\text{N}/^{13}\text{C}$ -labelled CFP-10 and ESAT-6 were prepared as described previously (Renshaw *et al*,

2002, 2004). In addition, $^{13}\text{C}/^1\text{H}$ HMQC-NOESY spectra were acquired from samples of the complex in which only the nonaromatic residues were uniformly $^{15}\text{N}/^{13}\text{C}$ labelled. This was achieved by the preparation of both proteins from *Escherichia coli* grown in labelled minimal media supplemented with 50 mg/l of L-histidine, L-tyrosine, L-phenylalanine and L-tryptophan (Carr *et al*, 2003). The mixed complexes of labelled CFP-10 bound to nonlabelled ESAT-6 and *vice versa* were produced by mixing equimolar solutions of the purified proteins at room temperature in 25 mM NaH_2PO_4 , 100 mM NaCl and 0.02% (w/v) NaN_3 , pH 6.5, with the individual proteins at a concentration of 5–15 μM . The complex was concentrated by ultrafiltration to give 0.35 ml NMR samples containing 0.9–1.5 mM CFP-10·ESAT-6 complex in either a 90% $\text{H}_2\text{O}/10\%$ D_2O or 100% D_2O buffer as appropriate.

Protein corresponding to a truncated variant of CFP-10 lacking the final 14 C-terminal residues (Asp87–Phe100) was prepared from a pET28a-based *E. coli* expression vector, which was produced using a PCR-based approach, essentially as described previously (Renshaw *et al*, 2002). Purification of the expressed protein was carried out in two stages using a 10 ml Q-Sepharose column (Renshaw *et al*, 2002), with truncated CFP-10 eluted from the column in the 50 mM NaCl step at pH 8.0 and in the 20 mM NaCl step at pH 5.8.

C-terminally truncated ESAT-6 corresponding to residues 1–84 is produced as a by-product during purification of the full-length protein. The two species are separated by anion exchange (Renshaw *et al*, 2002), with the truncated species eluted from the column in the 100 mM NaCl step.

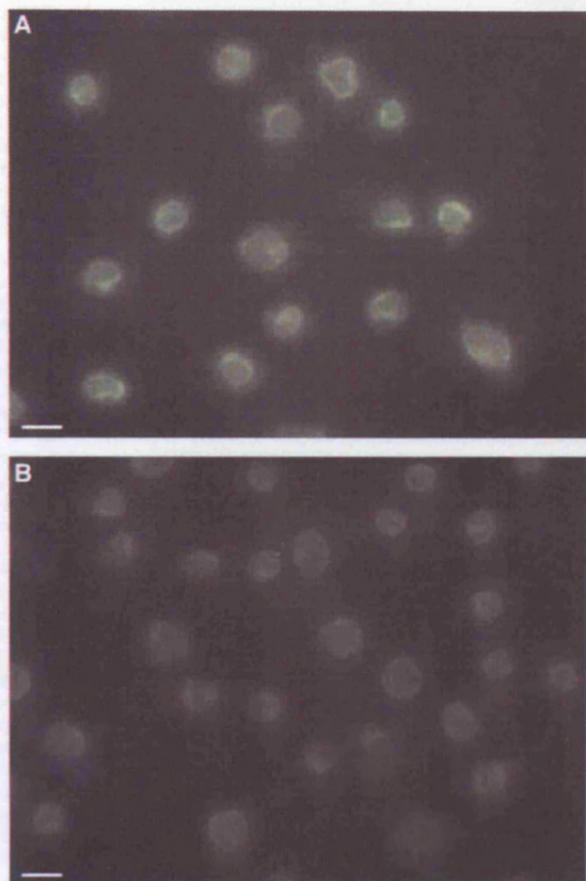


Figure 5 Blocking of fluorescently labelled CFP-10·ESAT-6 complex binding to U937 cells by competition with unlabelled complex. (A) Typical fluorescence images obtained for U937 cells after incubating with 1 μ M Alexa Fluor 546-labelled CFP-10·ESAT-6 complex for 15 min at 4°C. (B) The dramatic effect of adding a 20-fold molar excess of unlabelled complex. Clearly, the fluorescence labelling of U937 cells in the presence of an excess of unlabelled CFP-10·ESAT-6 complex is several orders of magnitude lower, which indicates that binding is mediated via the protein complex and not the attached fluorophore. The size bars shown correspond to 10 μ m.

NMR spectroscopy

NMR spectra were acquired at 35°C on either an 800 MHz Varian Inova or a 600 MHz Bruker Avance spectrometer. The 2D and 3D spectra recorded to obtain essentially complete sequence-specific backbone and side-chain assignments for CFP-10 and ESAT-6 in the complex, and to obtain conformational constraints for structural calculations were as follows: ^1H TOCSY and NOESY; $^{15}\text{N}/^1\text{H}$ HSQC, TOCSY-HSQC and NOESY-HSQC; $^{13}\text{C}/^1\text{H}$ HCCH-TOCSY and HMQC-NOESY; and $^{15}\text{N}/^{13}\text{C}/^1\text{H}$ HNCACB, CBCA(CO)NH and HBHA(CB-CACO)NH, as described previously (Renshaw *et al*, 2004).

The 3D NMR data were processed using NMRPipe (Delaglio *et al*, 1995), with linear prediction used to extend the effective acquisition times by up to 1.5-fold in F_1 and F_2 and mild resolution enhancement applied in all dimensions using a shifted sine-squared function. Apart from the omission of linear prediction, the 2D spectra were similarly processed using Varian or Bruker software. All the spectra were analysed using the program XEASY (Bartels *et al*, 1995).

Structural calculations

The family of converged CFP-10·ESAT-6 complex structures was calculated in a two-stage process using the program

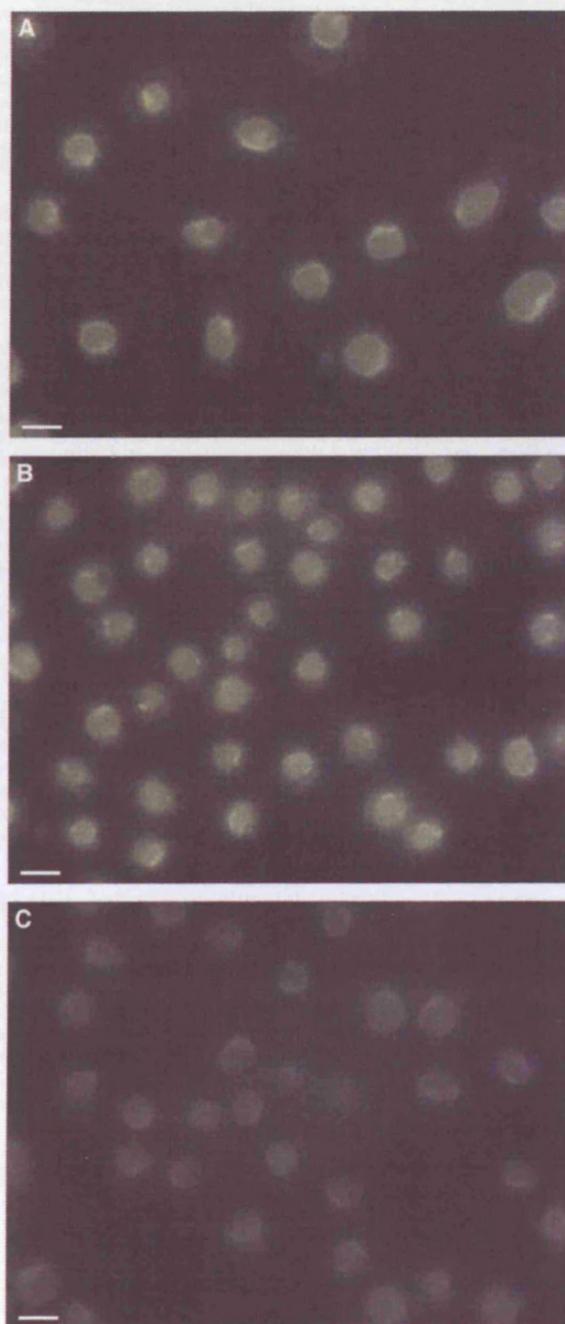


Figure 6 Dramatically reduced binding of the CFP-10·ESAT-6 complex containing C-terminally truncated CFP-10 to U937 cells. Panel (A) illustrates the typical fluorescence localisation observed for U937 cells incubated with full-length labelled complex, while panels (B) and (C) show comparable images obtained with labelled complexes containing C-terminally truncated ESAT-6 and CFP-10, respectively. Removal of the exposed C-terminus of ESAT-6 appears to have no significant effect on the ability of the complex to bind to U937 cells. In contrast, the fluorescence associated with U937 cells incubated with the complex containing C-terminally truncated CFP-10 is several orders of magnitude lower than that observed for the intact complex, which indicates that the flexible C-terminus of CFP-10 forms an essential component of the binding site for the cell surface receptor. The size bars shown correspond to 10 μ m.

CYANA (Herrmann *et al*, 2002). Initially, the combined automated NOE assignment and structure determination protocol (CANDID) was used to automatically assign the NOE crosspeaks identified in 3D ¹⁵N- and ¹³C-edited NOESY spectra of the complex and to produce preliminary structures. Subsequently, several cycles of simulated annealing combined with redundant dihedral angle constraints (REDAC) to increase convergence were used to produce the final converged CFP-10·ESAT-6 complex structures (Muskett *et al*, 1998; Lemercinier *et al*, 2001; Carr *et al*, 2003). The input for the CANDID stage primarily consisted of essentially complete ¹⁵N, ¹³C and ¹H resonance assignments for the nonexchangeable groups in the CFP-10·ESAT-6 complex and four manually picked NOE peak lists obtained from 3D ¹⁵N- and ¹³C-edited NOESY spectra of complexes in which only one protein was labelled. In the ¹⁵N-edited spectra, 1165 NOE peaks were identified with labelled CFP-10 and 1237 with labelled ESAT-6, and in the ¹³C-edited spectra 1962 NOEs with CFP-10 and 2580 with ESAT-6 were identified. In addition, the CANDID stage included ϕ and ϕ dihedral angle constraints for 95 residues in CFP-10 and 89 in ESAT-6, which were obtained from the ¹³C and ¹H chemical shifts of backbone resonances using TALOS (Cornilescu *et al*, 1999). The CANDID calculations were carried out using the default parameter settings in CYANA 1.0.6 apart from slightly increasing the chemical shift tolerances to 0.03 ppm for ¹H and 0.4 ppm for ¹⁵N and ¹³C.

The final converged CFP-10·ESAT-6 complex structures were produced from 100 random starting conformations using a torsion angle-based simulated annealing protocol combined with six cycles of REDAC (Muskett *et al*, 1998; Lemercinier *et al*, 2001; Carr *et al*, 2003). The calculations were mainly based on 3315 nonredundant, NOE-derived upper distance limits, assigned to unique pairs of protons using CANDID and corresponding to over 90% of the NOE peaks identified. However, constraints were also included for ϕ and ϕ dihedral angles in 184 residues and for hydrogen bonds formed by 37 residues with slowly exchanging backbone amide signals and where the hydrogen bond acceptor was unambiguous in preliminary structures (residues 22–26, 28–33, 58, 59, 61, 62, 65, 66, 68, 69 and 72 in CFP-10, and 28–33, 35, 36, 39, 62, 63 and 65–70 in ESAT-6). Slowly exchanging backbone amides in the complex were identified from a series of ¹⁵N/¹H HSQC spectra recorded over a period of several hours after dissolving samples of the complex in D₂O. The final family of CFP-10·ESAT-6 complex structures obtained were analysed using the programs CYANA, PROCHECK and MOLMOL, which included standard combined distance and orientation-based searches for hydrogen bonds and salt bridges (Koradi *et al*, 1996; Laskowski *et al*, 1996; Herrmann *et al*, 2002). Coordinates for the family of converged CFP-10·ESAT-6 complex structures, together with the NMR constraints, have been deposited in the Protein Data Bank under accession number 1wa8.

References

- Alderson MR, Bement T, Day CH, Zhu L, Molesh D, Skeiky YAW, Coler RN, Lewinson DM, Reed SG, Dillon DC (2000) Expression cloning of an immunodominant family of *Mycobacterium tuberculosis* antigens using human CD4⁺ T cells. *J Exp Med* 191: 551–559
- Bartels C, Xia T-H, Billeter M, Güntert P, Wüthrich K (1995) The program XEASY for computer-supported nmr spectral analysis of biological macromolecules. *J Biomol NMR* 5: 1–10
- Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284: 1520–1523
- Berthet F-X, Ramussen PB, Rosenkrands I, Andersen P, Gicquel B (1998) A *Mycobacterium tuberculosis* operon encoding ESAT-6 and a novel low-molecular-mass culture filtrate protein (CFP-10). *Microbiology* 144: 3195–3203
- Brosch R, Pym AS, Gordon SV, Cole ST (2001) The evolution of mycobacterial pathogenicity: clues from comparative genomics. *Trends Microbiol* 9: 452–458
- Carr MD, Bloemink MJ, Dentten E, Whelan AO, Gordon SV, Kelly G, Frenkiel TA, Hewinson RG, Williamson RA (2003) Solution structure of the *Mycobacterium tuberculosis* complex protein

Fluorescence microscopy

Samples of complexes corresponding to full-length CFP-10 bound to full-length ESAT-6, truncated CFP-10 bound to full-length ESAT-6 and full-length CFP-10 bound to truncated ESAT-6 were labelled with the fluorophore Alexa Fluor 546 (Molecular Probes) by incubating a 10-fold molar excess of the succinimidyl ester derivative of the dye with the respective complexes in a 25 mM NaH₂PO₄ and 100 mM NaCl, pH 7.5, buffer at room temperature overnight. At pH 7.5, the reactive succinimidyl ester group on the fluorophore is able to react with the N-terminal amino group of the two proteins, but not with charged lysine side-chain amino groups. Excess dye was removed by dialysis and the extent of labelling (typically 1.5–1.9:1) determined from the absorbance of the labelled complex at 280 and 556 nm, as per the supplier's instructions.

Primary monocyte, monocyte-derived macrophages, NIH-3T3 and COS-1 cells were grown directly on glass coverslips in appropriate media. The MonoMac 6 and U937 monocyte cell lines were initially grown in suspension and then allowed to adhere to glass coverslips precoated with 160 µg/ml poly-L-lysine for 20 min at 37°C. To assay for potential binding of the full-length CFP-10·ESAT-6 complex to the surface of specific cell types, cells adhered to coverslips were incubated with 1 µM Alexa Fluor 546-labelled complex for 15 min in PBS at either room temperature or 4°C. Nonbound complex was removed by two PBS washes prior to fixing of the cells with 4% (w/v) paraformaldehyde and permeabilisation with 0.2% (v/v) Triton X-100. The coverslips were mounted onto slides using ProLong antifade reagent (Molecular Probes) and stored at room temperature in the dark until dry. Fluorescence microscopy was carried out using a Nikon TE300 inverted microscope and the images recorded with a Hamamatsu CCD camera.

Similarly, U937 monocyte cells were incubated with 1 µM samples of Alexa Fluor 546-labelled combinations of truncated and full-length complexes for 15 min at 4°C, to minimise cell wall fluidity and possible receptor cycling, prior to being washed, fixed and imaged as described above. The blocking experiments were also carried out with U937 cells, which were incubated with a solution containing 1 µM labelled full-length complex and a 20-fold molar excess of unlabelled complex for 15 min at 4°C.

Acknowledgements

This work was initially supported by the award of a PhD studentship to Philip Renshaw from the Biotechnology and Biological Sciences Research Council and the Veterinary Laboratories Agency. Recent support has been provided by a project grant from the Wellcome Trust (066047). Kirsty Lightbody is supported by a PhD studentship from the Department for Environment, Food and Rural Affairs. Mark Carr is a member of the *Mycobacterium tuberculosis* Structural Genomics Consortium.

- MPB70. From tuberculosis pathogenesis to inherited human corneal disease. *J Biol Chem* 278: 43736–43743
- Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honoré N, Garnier T, Churcher C, Harris D, Mungall K, Basham D, Brown D, Chillingworth T, Connor R, Davies RM, Devlin K, Duthoy S, Feltwell T, Fraser A, Hamlin N, Holyroyd S, Hornsby T, Jagels K, Lacroix C, McLean J, Moule S, Murphy L, Oliver K, Quail MA, Rajandream M-A, Rutherford KM, Rutter S, Seeger K, Simon S, Simmonds M, Skelton J, Squares R, Squares S, Stevens K, Taylor K, Whitehead S, Woodward JR, Barrell BG (2001) Massive gene decay in the leprosy bacillus. *Nature* 409: 1007–1011
- Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J Biomol NMR* 13: 289–302
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6: 277–293
- Gey van Pittius NC, Gamielien J, Hide W, Brown GD, Siezen RJ, Beyers AD (2001) The ESAT-6 gene cluster of *Mycobacterium tuberculosis* and other high G+C Gram-positive bacteria. *Genome Biol* 2:Research. 0044.1–0044.18

- Guinn KM, Hickey MJ, Mathur SK, Zakel KL, Grotzke JE, Lewinsohn DM, Smith S, Sherman DR (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of *Mycobacterium tuberculosis*. *Mol Microbiol* **51**: 359–370
- Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* **319**: 209–227
- Hsu T, Hinigley-Wilson SM, Chen B, Chen M, Dai AZ, Morin PM, Marks CB, Padiyar J, Goulding C, Gingery M, Eisenberg D, Russell RG, Derrick SC, Collins FM, Morris SL, King CH, Jacobs WR (2003) The primary mechanism of attenuation of *Bacillus Calmette-Guérin* is a loss of secreted lytic function required for invasion of lung interstitial tissue. *Proc Natl Acad Sci USA* **100**: 12420–12425
- Koradi R, Billeter M, Wüthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* **14**: 51–55
- Kwiatkowska K, Sobota A (1999) Tyrosine phosphorylation/dephosphorylation controls capping of Fcγ receptor II in U937 Cells. *Cell Motil Cytoskeleton* **42**: 298–314
- Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**: 477–486
- Lemercinier X, Muskett FW, Cheeseman B, McIntosh PB, Thim L, Carr MD (2001) High resolution solution structure of human intestinal trefoil factor and functional insights from detailed structural comparisons with the other members of the trefoil family of cell motility factors. *Biochemistry* **40**: 9552–9559
- Lightbody KL, Renshaw PS, Collins ML, Wright RL, Hunt DM, Gordon SV, Hewinson RG, Buxton RS, Williamson RA, Carr MD (2004) Characterisation of complex formation between members of the *Mycobacterium tuberculosis* complex CFP-10/ESAT-6 protein family; towards an understanding of the rules governing complex formation and thereby functional flexibility. *FEMS Microbiol Lett* **238**: 255–262
- Mahairas GG, Sabo PJ, Hickey MJ, Sing DC, Stover CK (1996) Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent *M. bovis*. *J Bacteriol* **178**: 1274–1282
- Muskett FW, Frenkiel TA, Feeney J, Freedman RB, Carr MD, Williamson RA (1998) High resolution structure of the N-terminal domain of tissue inhibitor of metalloproteinases-2 and characterisation of its interaction site with matrix metalloproteinase-3. *J Biol Chem* **273**: 21736–21743
- Pym AS, Brodin P, Brosch R, Huerre M, Cole ST (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines *Mycobacterium bovis* BCG and *Mycobacterium microti*. *Mol Microbiol* **46**: 709–717
- Pym AS, Brodin P, Majlessi L, Brosch R, Demangel C, Williams A, Griffiths KE, Marchal G, Leclerc C, Cole ST (2003) Recombinant BCG exporting ESAT-6 confers enhanced protection against tuberculosis. *Nat Med* **9**: 533–539
- Renshaw PS, Panagiotidou P, Whelan A, Gordon SV, Hewinson RG, Williamson RA, Carr MD (2002) Conclusive evidence that the major T-cell antigens of the *Mycobacterium tuberculosis* complex ESAT-6 and CFP-10 form a tight, 1:1 complex and characterisation of the structural properties of ESAT-6, CFP-10 and the ESAT-6.CFP-10 complex: implications for pathogenesis and virulence. *J Biol Chem* **277**: 21598–21603
- Renshaw PS, Veverka V, Kelly G, Frenkiel TA, Williamson RA, Gordon SV, Hewinson RG, Carr MD (2004) Letter to the editor: sequence-specific assignment and secondary structure determination of the 195-residue complex formed by the *Mycobacterium tuberculosis* proteins CFP-10 and ESAT-6. *J Biomol NMR* **30**: 225–226
- Rosenkrands I, Weldingh K, Jacobsen S, Hansen CV, Florio W, Gianetri I, Andersen P (2000) Mapping and identification of *Mycobacterium tuberculosis* proteins by two-dimensional gel electrophoresis, microsequencing and immunodetection. *Electrophoresis* **21**: 935–948
- Sasseti CM, Rubin EJ (2003) Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci USA* **100**: 12989–12994
- Schnappinger D, Ehrh S, Voskuil MI, Liu Y, Mangan JA, Monahan IM, Dolganov G, Efron B, Butcher PD, Nathan C, Schoolnik GK (2003) Transcriptional adaptation of *Mycobacterium tuberculosis* within macrophages: insights into the phagosomal environment. *J Exp Med* **198**: 693–704
- Skjöt RLV, Brock I, Arend SM, Munk ME, Theisen M, Ottenhoff THM, Andersen P (2002) Epitope mapping of the immunodominant antigen TB10.4 and the two homologous proteins TB10.3 and TB12.9, which constitute a subfamily of the *esat-6* gene family. *Infect Immun* **70**: 5446–5453
- Skjöt RLV, Oettinger T, Rosenkrands I, Ravn P, Brock I, Jacobsen S, Andersen P (2000) Comparative evaluation of low-molecular-mass proteins from *Mycobacterium tuberculosis* identifies members of the ESAT-6 family as immunodominant T-cell antigens. *Infect Immun* **68**: 214–220
- Stanley SA, Raghaven S, Hwang WH, Cox JS (2003) Acute infection and macrophage subversion by *Mycobacterium tuberculosis* require a specialised secretion system. *Proc Natl Acad Sci USA* **100**: 13001–13006
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**: 4876–4882
- Volkman HE, Clay H, Beery D, Chang JCW, Sherman DR, Ramakrishnan L (2004) Tuberculous granuloma formation is enhanced by a *Mycobacterium* virulence determinant. *PLoS Biol* **2**: 1946–1956
- Wards BJ, de Lisle GW, Collins DM (2000) An *esat6* knockout mutant of *Mycobacterium bovis* produced by recombination will contribute to the development of a live tuberculosis vaccine. *Tubercle Lung Dis* **80**: 185–189
- World Health Organisation (WHO) Geneva (2004) Global tuberculosis control: WHO Report 2004. WHO/HTM/TB/2004.331